

# Open Research Online

---

The Open University's repository of research publications and other research outputs

## A case study of Emirati females and an international EFL oral proficiency test: Does one size fit all?

### Thesis

#### How to cite:

AlHafidh, Gail Karen (2011). A case study of Emirati females and an international EFL oral proficiency test: Does one size fit all? EdD thesis The Open University.

For guidance on citations see [FAQs](#).

© 2011 The Author



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Version of Record

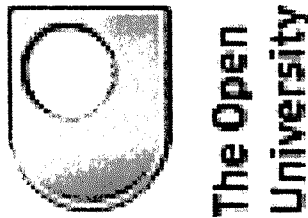
Link(s) to article on publisher's website:  
<http://dx.doi.org/doi:10.21954/ou.ro.0000d629>

---

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

---

[oro.open.ac.uk](http://oro.open.ac.uk)



**A Case Study of Emirati females  
and an international EFL oral proficiency test. Does  
one size really fit all?**

Doctorate of Education: 2011

Submitted by: Gail Karen AlHafidh B.A (Hons) M.A

Supervisors: Dr Barbara Heins & Dr Helen Peters

Date of Submission: 31 January 2011  
Date of Award: 10 July 2011

## **Abstract**

For many students, English as a foreign language (EFL) assessments are high-stake examinations, the results of which will determine their future study and career paths. This thesis will present data gathered from questionnaires of students and examiners, filmed interviews, audio recordings and focus group feedback. The key question posed is: is it possible for international EFL exams to have global applicability and therefore maintain test validity? Furthermore, to what extent should international EFL test writers take into account the regional, socio-cultural context of the recipient student body when making question choices and devising assessment criteria? This thesis attempts to address these questions through interpretive case study research of oral interview assessment in a female campus of a tertiary college in the United Arab Emirates (UAE). CAT (Communication Accommodation Theory) was the theoretical tool used to examine the interaction between examiners and candidates during the IELTS-style oral assessment.

The study looked at the questions typically asked in an international EFL interview and the reaction of the participants, both linguistically and behaviourally, to the context of the interview and the method of assessment of the communication in those interviews. The aim was to have a better understanding of how female Emirati candidates respond to the various aspects of an international EFL speaking assessment in the context of their own cultural, social and religious constructs.

Analysis of the data reveals that there is a mismatch between the perceptions of the examiners and the participants, in several areas, and that this has the potential to affect grade outcomes, as seen in the case study interviews. The study concludes that there is scope for both a broader range in the choice of questions in speaking assessments and a need for examiners to be prepared to choose questions more judiciously, in line with the cultural context of the candidates and that this is possible without jeopardising the validity of the assessment. The findings also show that there are clear differences between the grading of face to face and audio

recorded interviews and that these should be considered when grading criteria are written. Overall the study contributes a variety of insights into the field of oral assessment and has implications for test writers, assessors, candidates and publishers, since, in the case of international EFL oral assessments, it appears that 'one-size' does not fit all.





## **IMAGING SERVICES NORTH**

Boston Spa, Wetherby  
West Yorkshire, LS23 7BQ  
[www.bl.uk](http://www.bl.uk)

# BLANK PAGE IN ORIGINAL

### Statement of originality

This research is both original in its content and its context. It is a case study of Emirati female students, in the context of their own college campus, undergoing the oral component of an EFL (English as a Foreign Language) test which, in this case, is based on the internationally recognized IELTS ((International English Language Testing System) exam. Authentic assessment grading criteria have been used from the publicly available IELTS website and assessment materials have been originally designed by the researcher in the model of existing IELTS practice materials, again publicly available on the official IELTS website. The student participants, the examiners, the focus group participants and the questionnaire respondents are all members of the same community, that is, the student and teaching body at the female campus of the college in Sharjah, United Arab Emirates. The questionnaires and the focus group stimulus questions are the original work of the researcher.

There is currently no published empirical data relating to Emirati female veiled/nonveiled students undertaking oral examinations, to my knowledge, and therefore this research is of great value in exposing the issues and concerns relating to EFL oral testing in this specific context. The researcher uses CAT (Communication Accommodation Theory) to underpin the analysis of the interaction occurring during the oral examinations. To date, CAT has been used largely in the Health and Law enforcement environments to analyse patient/doctor interactions and police/defendant interviews. This research is therefore also original in that it is broadening the scope and application of CAT to EFL oral testing.

The findings of this original research will be of direct relevance to EFL professionals, not only in the Gulf region but will also be of specific interest to anybody involved in testing speaking skills especially within other cultural contexts and with veiled/partially veiled candidates.

## Table of Contents

	<b>Page Number</b>
Abstract	2
Statement of Originality	5
Introduction	8
<b>CHAPTER 1: LITERATURE REVIEW</b>	<b>11</b>
Communication	11
Assessing Spoken communication	17
Communication Accommodation Theory: CAT	23
Non-verbal interaction	28
Current research: IELTS and IELTS-style assessments	39
<b>CHAPTER 2: METHODOLOGY</b>	<b>49</b>
Case Study	49
Pilot project	51
Main Case Study	68
CAT	84
<b>CHAPTER 3: DATA ANALYSIS</b>	<b>87</b>
Linguistic communication grading criteria	88
Face to face grading and application of CAT	100
Questionnaire Data	108
Focus Group Data	115
<b>CHAPTER 4: DISCUSSION AND CONCLUSIONS</b>	<b>120</b>
Research Questions and discussion of results	121
Research methodology reviewed	123
Significance of research findings	127
Conclusions	144
Opportunities for further research	149
References	155
Appendices	ii

## **Introduction**

Success in English as a foreign language (EFL) is routinely measured by one of the major providers in international EFL assessment, the American based TOEFL (Test of English as a Foreign Language), TOEIC (Test of English for International Communication) and the Australian/British IELTS ( International English Language Testing System). On the one hand, benchmarking to international standards means that students can be equipped with internationally-recognised qualifications that can be used to springboard their careers in the global job-market: according to the IELTS official website students taking IELTSs do so for a variety of reasons including: job opportunities, registration on further and higher education courses and immigration. On the other hand, there is a risk that students may be disadvantaged if they do not share the same values and ideological systems as those of the test writers (Wallace 1997; Edge 1994; Phillipson 1992; Pennycook 1994). Wallace, writing about the global implications of curriculum and materials design, highlighted the problem of IELTS writing tasks being prejudicial to the success of students who did not share the same schemata as the test writers and therefore could not write the tasks effectively, rendering the tests themselves unrealistic as tools of writing skills measurement. Edge also warned about the presumption of course and test writers that they shared values and ideological systems with the test takers. For students to be able to be assessed equitably, globally, test items must be equally accessible; it is crucial therefore that the students non-linguistic schemata are also considered and accommodated as far as possible (Carrell, Devine and Eskey 1988).

The rationale behind this research was therefore to study the extent to which international EFL tests, in this instance tests based on the IELTS model, can really be internationally applicable. The Emirati, female student participators in this research are studying towards the IELTS exam and, in preparation for this, encounter IELTS-style (in-house) assessments as part of their coursework. Their life experience and cultural and religious backgrounds are very different from many of the other students globally

who are also preparing for the exam. The research focuses on oral communication assessment and asks whether it is possible to have a “one size fits all” international E.F.L exam.

The starting point for this case study was an observation that students seemed to be having a mixed experience in their IELTS speaking exams, based on anecdotal evidence from the researcher’s own students. Students were commenting on the topics that they had to discuss and there was a definite hierarchy of perceived difficulty attributed to certain topics. This led to a questioning of the oral assessment process as a whole. Was the test an artificially constructed ‘performance’ in which students were asked to deliver linguistic items that were measurable (according to a set of pre-defined criteria) in order to be rated a particular Band? In practice sessions with the students, it became clear that students who were less gregarious often scored less well than expected, despite having the same linguistic ability as their ‘chatty’ peers. Why was this? Surely the assessment was meant to be a test of language skills not personality, of exchange of information, not an exchange of friendly repartee.

There were other factors that came to light too that students felt perhaps went in their favour or against them. One was the wearing of a veil. The students in the case study are all Emirati females, the context being an all female technical college. Some students felt obliged to remove their veils for the speaking exam fearing wearing them might jeopardize their chance for an optimum grade, others wore theirs for the opposite reason: perhaps they felt that wearing a veil would be to their advantage. Some students expressed worry that they did not speak ‘like an American’ or ‘like their English teacher’ and that would also count against them. Perceptions were that in order to attain a good grade in their speaking exam, there were strategies other than simply the linguistic competencies that would hold sway. This was particularly interesting because the grading criteria used at the college for speaking assessments, the IELTS publically-available band criteria, are purely based on features of linguistic ability.

I decided that this needed further investigation and that a case study approach was the best option, in this context. The research was to try to understand and interpret the world in terms of its actors (students and examiners) and as such would be interpretive and subjective.

The research data includes 11 video-recorded, mock in-house IELTS-style (International English Testing System) interviews with female candidates and both male and female examiners, all of whom are accredited IELTS examiners; questionnaire data from second year students preparing to take their IELTS exam and IELTS examiners, and, thirdly, audio-recorded focus group meetings between the researcher and the participants of the study. A Case Study approach was decided upon because the context of the research was contained and easily defined, yet the complexity within this specific context was of particular interest and generated the research questions. Stake (1995, p. xi) describes the case study as the “study of the particularity and complexity of a single case, coming to understand its activity within important circumstances. The research focuses on a specific context of an all-female student campus, where students are studying a variety of program choices but all are united in their goal, as a graduation requirement, of reaching a particular IELTS Band level. This single case study has unique complexities that can be seen to have relevance in the field of EFL assessment, specifically oral assessment and seeks to shed light on the notion that there can be a universal test of spoken English.

## **CHAPTER 1**

### **Literature Review**

This literature review will present the background to the case study research and will be divided into the following sections:

- 1. Communication. What is actually being assessed?
- 2. Assessing spoken communication.
- 3. Communication Accommodation Theory (CAT).
- 4. Nonverbal interaction
- 5. Current research: IELTS and IELTS-style assessments

#### **1.1 Communication**

Language, verbal language, is just one aspect of the human communication system which is defined by the Concise Oxford dictionary (ed Thompson, 1995) as 'the act of imparting information' or 'social intercourse'. These definitions of communication are very narrow for such a complex concept which will be explained in more detail below. Communication is an interactional, two-way process involving communicator and recipient in a multi-layered, complex interchange. Language is therefore only one of the many social practices that operate interactively to represent and make meaning in communication (Halliday 1978; Kress and Van Leeuwen 2001; Fairclough 1989). Norris (2004 ) identifies those social practices and has developed the term 'multi-modal interaction' incorporating gesture, gaze , pose, material and environmental surroundings, individual perceptions and experiences as all underpinning language and being central to understanding the process of communication. Furthermore, language carries with it social, cultural and ideological meanings and associations which are perceived differently by individuals, especially where cultures interface. Lustig & Koester (2006. p.10) also add the cultural dimension, implicit in Norris's reference to individual perceptions, to their definition of communication:

Communication is a symbolic, interpretive, transactional, contextual process in which people create shared meanings.

They define these four features of spoken communication. Each aspect of the communication process, it is suggested, is part of a core set of basic characteristics central to all communication, including intercultural interactions. Here the symbolic feature is the expression of body language and gesture whilst the interpretive refers to the negotiation of understanding and perceptions of understanding. The transactional aspect refers to the ebb and flow of messages in communication and the contextual to the physical, social and interpersonal setting. All these can be seen as the broadly accepted aspects of communication. The intersection of cultures presents a new dimension to the process in that the interpretation of symbols, for example, may not be the same for each participant and the mutual influence of both interpretations may lead to a very different understanding of the transaction, rather than Lustig and Koester's 'shared meaning'.

### Culture

At this point, it is necessary to define what is meant by the term *culture* in the context of intercultural communication. Lustig and Koester (2006. p.25) define culture as:

a learned set of shared interpretations about beliefs, values, norms, and social practices, which affect the behaviours of a relatively large group of people.

Kramsch (1998 p.10) links her definition to the nature/nurture debate and questions whether culture is taught through socialization or a given at birth. She states that:

culture can be defined as membership in a discourse community that shares a common social space and history, and common imaginings.



There are many such definitions that vary according to the discipline involved but for the purposes of this study, the following definition will be used as it seems to incorporate the sense of shared perceptions and experiences mentioned already but also includes the notion that these values not only have a historical significance but also a future importance in helping to shape and define the members of a particular social group:

Culture: the system of shared beliefs, values, customs, behaviours, and artifacts that the members of society use to cope with their world and with one another, and that are transmitted from generation to generation through learning.

(<http://www.umanitoba.ca/faculties/arts/anthropology/courses/122/module1/culture.html>)

It follows that intercultural communication involves the interface of cultures as part of Lustig and Koester's (2006) *symbolic, interpretive, transactional* and *contextual* process. It involves a convergence of ways of thinking, speaking and viewing the world and has implications for a variety of disciplines interested in human interaction, not least the field of linguistics and language teaching, and the teaching and assessment of speaking skills. It is, as Kramsch (1993 p.2) writes 'a small miracle' when understanding and shared meaning, across cultures, occurs and it requires an acceptance of complexity and ambiguity rather than the view that language teaching is 'the teaching of forms to express universal meaning'. When native speakers speak to each other, they are more likely to understand the humour, the idioms and metaphors that shape their everyday communication and underpin their linguistic exchange. However, the parameters of cultural boundaries, even within a native speaker to native speaker context, are dynamic and flexible such that an American speaking to a British person still may not be able to share a joke or understand a particular idiom. Non-native speakers, even when linguistically competent may find it very difficult to bridge this gap because of the extent of their lack of cultural

knowledge. This research will explore to what extent this lack of cultural knowledge may be a hindrance in EFL oral assessments.

There are various concepts underpinning the theoretical framework that need to be clarified. Firstly, *cultural identity* and *intercultural communication* specifically in terms of the participants involved in the research, both students and assessors, are core concepts throughout. Imahori and Lanigan, (1989) when working towards creating a model of intercultural communication competence, identified knowledge, motivation and skills as key components affecting the outcome of an interaction. They argued that cultural identity manifests itself via a knowledge base of communication rules which are linguistic rules that are also either culture specific or general. Thus, speakers from a particular culture know what language to use and when, how and why to use it. Furthermore, speakers have an attitude or opinion about the other's culture, particularly in the case of intercultural communication, and via their communication skills they express themselves to a greater or lesser degree, with empathy, accommodation, respect and so on. Kramsch (1998) agrees that culture is reflected and shaped by the linguistic choices we make and, that in intercultural exchanges, both speaker and hearer jointly construct the meaning of an utterance to reach mutual understanding.

Cultural identity is then not simply about presenting particular national traits in both character and appearance, it is much more than that; it is a dynamic phenomenon that manifests itself in many ways but through intercultural communication it is evident in the way interlocutors jointly construct meaning using the linguistic skills, motivation and knowledge identified above by Imahori and Lanigan (1989).

According to Halliday,

“language neither drives culture nor is driven by it; the old question about which determines which can be set aside as irrelevant, because the relation is not one of cause and effect but rather one of

realization; that is culture and language co-evolve in the same relationship as that in which, within language, meaning and expression co-evolve.” (Halliday, 1993. p11).

Any test of language must therefore also acknowledge and be sensitive to the cultural context in which it is being administered.

#### Native speaker competence

Another two concepts central to the research are that of *native speaker* and that of *World English*. The term ‘native speaker competence’ is commonly understood to mean that the speaker speaks fluently and accurately, as though the language spoken is their mother tongue. The concept might at first indicate that you can either be a native speaker or not, and that it is an absolute condition. This is not the case as far as assessment in EFL is concerned. The term ‘native speaker competence’ exists on a continuum with the beginner language learner at one end of the continuum and ‘native speaker’ competency at the other. In published IELTS Banding assessment criteria, to achieve a Band 9 (the highest in a scale 1-9) in the category of Grammatical Range and Accuracy, the candidate must: “produce consistently accurate structures apart from ‘slips’ characteristic of native speaker speech” (appendix 1). This descriptor is problematic on several levels: not only is the examiner asked to discern what constitutes ‘structural accuracy’ (this may vary between American and British English, for example) but also to make a judgment call on what ‘slips’ are characteristic of ‘native speaker’ speech. Are, for example, confusing ‘stationery’ with ‘stationary’ or ‘complimentary’ with ‘complementary’ slips characteristic of native speech? Or, are slips of subject-verb agreement such as ‘the binding of the books were damaged’ characteristic of native speech? These are common errors made by some British, mother-tongue English speakers, so, would the person making this kind of error as a native speaker, be considered worthy of a Band 9?

The problem is further compounded by the fact that not all international EFL examiners are mother-tongue, first language English speakers so this kind of judgement call is potentially difficult to make. At the lower end of the scale in the Pronunciation category, to be awarded a Band 4, a candidate must “produce some acceptable features of English pronunciation but overall control is limited and there can be some strain for the listener”. Again there are problems with this descriptor in that assumptions are made about the fact that there are universally recognizable “features of English pronunciation” and that if these are not adhered to, some ‘strain’ is experienced by the listener. A person living in Southern England, a native speaker of English, may experience severe strain listening to a person living in Liverpool who is also a native speaker, and vice versa. So, without going into the finer descriptors of language, dialect and accent, what exactly is “English” pronunciation?

Melchers and Shaw (2003) have charted the roots of English from Anglo-Saxon times and note that linguistic variation and change occurred as speakers came into contact with other languages or dialects throughout the centuries, indicating that language is dynamic and evolves. Today, globalization has speeded up the rate at which language is evolving, in particular the English language, and the subsequent term ‘World Englishes’ has been coined to reflect the range and scope of English usage around the globe. In this research project, I hope to be able to shed light upon the extent to which the existence of ‘World English’ is recognised and reflected in the assessment criteria of international English exams. The dilemma today for examiners is that what used to be known as Standard English, whether British or American (despite differences in pronunciation, spelling, vocabulary and slightly in grammar) no longer exists. There are many forms of the English language that are not ‘standard’ American, British or Australian English. Sartor and Heng (2008) in defending the existence and importance of various forms of English go further to suggest that there exists a ‘form of language genocide’ they refer to as linguisticism, whereby other varieties of English are considered less valid. Reporting on trends in mainland China, they observe that Standard American or British

English are perceived as the main measures of competency in English language. According to Trudgill and Hannah (1994), Standard English is actually a dialect not an accent and there is no International Standard yet. In other words, as Trudgill and Hannah (1994) argue, a person could speak a variety of Standard English very well but with a Chinese accent, for example. Or, conversely, a person could have a very British accent but speak a non-British version of English as do some Indian speakers. World English therefore, in this research project, will be used as an umbrella term encompassing all the varieties of English spoken around the globe, including those varieties spoken by the examiners in EFL examinations. This broad range of varieties of English, make the assessment of spoken English particularly problematic.

### 1.2 Assessing spoken communication

Speech, like the written word, is contextual. It depends upon the context, the power relationship between the speakers, the schemata they share or differ in and, the topic being discussed. It can be rehearsed and structured, compliant with social norms and etiquette or, as is more common, dynamic, spontaneous and unpredictable. Unless the interaction has been rehearsed specifically, spoken communication usually involves a negotiation of meaning and for understanding to take place, the participants need to share understanding of those meanings using linguistic content as one tool to achieve this. Strategies employed for clarification, such as asking for repetition and rephrasing, facilitate that sharing of meaning.

Language has a purpose, whether its function is to ask for directions, explain a procedure or to tell someone a story (Gibbon, 1993). Assessing spoken language cannot therefore be simply about linguistic knowledge of the language, its grammar and lexis but must include its pragmatic use in a given context as one part of a complete communication strategy. This makes assessment of speech much more complex than measuring language proficiency as a single point on a linear scale using an objective test

instrument (Underhill, 2003). Underhill makes the point that historically, language tests, particularly oral tests have been too focused on the test instrument itself and not the human participants. Tests have been artificially constructed to 'test' language skills rather than assess linguistic or rather communicative ability. However, it is difficult to see how summative oral competency tests can practically be operated in any other way. Assessors need a means of measuring linguistic skill and the most practical way appears to be by using a sliding scale of criterion-based evidence of grammar, lexis, pronunciation and general fluency. Madsen (1983) reviewed 74 exams and found 84% measured grammar, 71% fluency, 67% vocabulary, 66% pronunciation, 63% 'appropriateness' and 37% 'other'. These 'other' factors are not clearly defined but Madsen points to listening comprehension, correct tone, reasoning ability, and the ability to ask for and understand clarification. So, is it an unachievable goal perhaps to try to assess speaking in any other way than has been done to date?

Fulcher (2003) argues that the starting point has to be 'defining the construct', in other words defining what it means to 'speak' and agreeing on what the observable and measurable features of speech are so that these can inform the design, rubrics and delivery, assessment tasks and the means of measuring the achievement of those tasks. He draws a distinction between competence and performance as 'inter' and 'intra' individual respectively such that communicative competence is an internalized, personal phenomenon (an individual's knowledge of vocabulary and grammatical rules, for example) as compared with communicative performance which is reliant on interaction. The two features of communication though are not mutually-exclusive; they are interdependent. Just as in any form of knowledge, knowing does not necessarily manifest itself in a tangible, measurable way. I can know how to make a cake, know what ingredients are required, the process of putting all those together but it does not mean I will necessarily succeed in making a cake at a given time and in a particular context. Competency, in this case, is not an indicator of performance.

### Test design

Returning now to language test design, the participants in those tests, the test-takers and the assessors should feel that the test is relevant and meaningful. The test constructs, should be guided by test purpose (Fulcher, 2003). We still have a dilemma, however. To a certain extent, the assessing of reading and listening is much more straight-forward than assessing speaking or writing. Either the test-taker hears correctly and can answer the questions, and the reader locates the information required to answer the question, or s/he does not. Listening and Reading assessments tend to be multiple choice or short-answers that have very few alternative answers possible. Free writing, on the other hand, is slightly more difficult to assess as there is much more scope for the candidate to interpret the question and respond at length. The grammar and punctuation, lexis and organization of the writing can all be assessed quite scientifically against pre-set criteria but the writer's interpretation of the question and style of response may vary and so there may be many more possible correct/incorrect responses and much more differentiation between grades. This is gross over-simplification of course, but generally speaking the assessment of listening and reading are much more dependent on the 'inter' individual, or competence, than writing and to a greater extent, speaking. Speaking relies on interaction (Goffman, 1981, Goodwin, 1990, Philips, 1972). Even Hamlet's famous soliloquy, 'to be or not to be' can be seen as an interaction with himself, with his alter ego.

### The nature of speech

Speech interaction in a test situation is not only about Fulcher's 'competence' but it is also about personality, mood, 'clicking' with who you are talking to, context (physical and cultural) and subject matter; speech interaction is a performance. Hamlet aside, speech generally involves a minimum of two parties and so the 'speech act' can be seen to be the sum of the two parties' contributions, in other words, speech is co-constructed. In Kramsch's (1986) model of Interactional Competence Theory (ICT) this is recognized and so the assessment of a candidate in an EFL oral assessment

becomes more about the interactional competence of the co-constructed communication (between assessor and candidate) than the individual performance. This leads to more problems for test creators; if speech is co-constructed, how can we assess a candidate when to a certain extent their production is reliant on the input of the assessor?

The speech act itself cannot be controlled if it is to replicate an authentic exchange, as far as possible. What can be controlled are the test rubrics and test content. IELTS tests are very restricted in terms of assessor input. Each question is scripted and examiners are asked to follow the script closely to maintain test validity internationally and to minimize the risk of examiner over or under input. The examiner, as interlocutor, plays a key role in the interview assessment.

### Role of the assessor

Nakatsuhara (2008) studied two interviewers with the same candidate and then had 22 raters independently grade the candidate's performance. The purpose was to try to ascertain the level of support given by an assessor and the influence that had on the candidate's grade outcome. The rubrics were not strictly controlled, the examiners were only provided with a set of questions and a photo. The results showed differences in grades for fluency and pronunciation and clearly indicated that the extent to which an assessor develops topics and reacts to the candidate's responses, in fact all examiner behavior, has an impact on the language produced by the candidate.

Interestingly, one of the examiners was an IELTS trained assessor and the candidate scored a lower grade in this case. Nakatsuhara (2008) suggests because of the standardization training that IELTS examiners undergo, this should not have occurred. The implication is that assessor training can reduce the difference in assessor behaviour, but to what extent does this then invalidate or 'corrupt' the communication that takes place? If the input from the assessor is scripted and, 'rehearsed', to what extent can the candidate's input, as co-constructor of the communication, be assessable? Their response is effectively to a scripted speech rather than a 'live' input with all its spontaneity and unpredictability.



### Assessor training

One of the aims of assessor training is to minimize rater differences so that any candidate at any time is likely to be graded consistently and fairly and to ensure that examiners are applying the grading criteria in the same way.

Mullen (1978) concluded that to be entirely reliable, in speaking assessments, two assessors are better than one because each assessor may rate differently for specific criteria (grammar or lexis for example) this being masked by a similar global grade. The evidence further suggests that where assessors are trained examiners, the reliability is higher (Shohamy, 1983. Morrison and Lee (1985) as cited in Fulcher 2003).

Lumley and Brown's research (1996) looked at how interviewers behaved and found that those who took on the role of the interviewer rather than that of a supportive teacher offering vocabulary support and rephrasing questions, were perceived as more difficult interviewers by the candidates. Brown and Hill (2007) also studied interviewer style and its effect on grade outcome in IELTS interview assessments. The results indicated that 'easier' (more lenient) interviewers changed topics more often and asked more direct, simpler questions whereas the interviewers perceived as more difficult asked more challenging questions. The research was conducted with trained IELTS examiners and seems to suggest that despite this training, examiners still offer differing amounts of support (behavioural) to candidates. One measure was of the number and length of turns for each interviewer. The easiest interviewer had the longest turns and used feedback more frequently. They also asked more factual type questions and fewer speculative questions. The easiest interviewer also made nine topic shifts compared with two of another interviewer. Brown and Hill's study concludes that 'there is no doubt that candidates can be disadvantaged or advantaged...in interviewer allocation' (p.55).

O' Loughlin (2007) explored the supposition that gender differences have a bearing on assessor behaviour. In his studies into the IELTS oral test, he found, contrary to previous research (Coates 1993, Tannen 1990) that gender did not have a significant impact. O'Sullivan (2000) found that male

assessors, with Arab students tend to give higher scores in assessments and that the same was true of women interviewers of Japanese students (male or female). O'Sullivan and Lu (2002) analysed data from IELTS interviews to determine the impact of deviation from the interview script, but found no systematic deviation points and no therefore no significant difference in candidate language, implying that interlocutor frames could be less prescriptive and more flexibility could be allowed without affecting the validity and reliability of the test. Further work would be needed to extend this study to incorporate gender differences in candidate behaviour, and a gender mix, assessor to candidate. To single out gender as a possible factor in rater difference is only one part of the story. Research by Berwick and Ross (1992) suggests that where reliability may differ is if examiners accommodate to the candidate based on their own cultural background. There are many other variables that can have an impact on the dynamics of the conversation as suggested earlier; including mood, context, personality, cultural background, first language, schemata and, previous experience of an oral exam situation. The goal of any assessment must be to create the opportunity, through carefully constructed questions, to ensure that the candidates are able to showcase their language skills without corrupting the essence of the communication act itself, that it should be negotiated meaning. Brown (2007b) writes that perhaps to aim for consistent assessor behaviour is 'over ambitious' and that raters should be allowed 'their individuality and internal variability' (p.138). As Berwick and Ross (1992) showed, successful communication is not only about the transmission of discrete linguistic items but also about who you are and how well you are able to 'connect' psychologically with your interlocutor. As we have seen, the assessment of communication is not simply about spoken language but there are many other variables to consider.

Communication Accommodation Theory (CAT) is one tool with which to explore communication in general and reveal the extent to which factors other than linguistic co-construction can determine the quality of communication. CAT acknowledges that communication is not simply

about speech acts or linguistic behavior but also paralinguistic and non-linguistic moves between interactants.

### 1.3 Communication Accommodation Theory (CAT)

CAT has emerged from SAT (speech accommodation theory) as a result of the acknowledgement that communication is not just about speech and speech styles.

CAT is based on three general assumptions:

1. Communicative interactions are embedded in sociohistorical context.
2. Communication is about both exchanges of referential meaning and negotiation of personal and social identities.
3. Interactants achieve the informational and relational functions of communication by accommodating their communicative behaviour, through linguistic, paralinguistic, discursive, and non-linguistic moves, to their interlocutor's perceived individual and group characteristics.

(Gallois et al, 2005, p.137)

CAT can be used to investigate the links between language, context and identity. Moreover, it attempts to conceptualize communication in both subjective and objective terms: what are the motives and intentions behind speakers' conscious (or unconscious) linguistic choices? How do listeners perceive these choices and react to them (Gallois, Ogay, Giles. 2005)? The basic premise is that humans need approval and that in order to seem more attractive, they try to reduce any dissimilarities between each other.

Conversely, if they feel threatened, they may wish to distance themselves from one another. The convergence and divergence strategies presented in CAT are of direct significance in a foreign or second language teaching context since it could be argued that learners will be striving to converge with the native speaker in an attempt to seem more fluent whereas the recipient (teacher) may be either maintaining their speech patterns or even

diverging away consciously or unconsciously in an attempt to 'lure' the student towards the target language. There is, therefore, potential tension between the linguistic and psychological features of communication.

Thakerar et al. (1982) distinguished between linguistic accommodation and psychological accommodation arguing that while speakers' linguistic shifts could be observed as either converging or diverging, speakers may perceive the opposite to be true. This is also the case for recipients. Hornsey and Gallois (1998) used CAT to explore how convergence and divergence could be applied to either person-based or group-based intercultural communication in a study of Chinese and Australian speakers. They found that individuals may be psychologically motivated to adjust or accommodate their linguistic, speech-related behaviour to gain approval, pursue relationship development and maintenance objectives or attain greater communication efficiency.

In the case of intercultural communication, all three CAT assumptions are pertinent to the study of spoken interaction, not least the third including its reference to paralinguistic features of communication. Similarly, all three assumptions with their reference to cultural concepts such as 'the sociohistorical context' and 'the negotiation of personal and social identities' offer a means of exploring the complexity of intercultural communication. CAT is an attempt to embrace all aspects of communication and in that goes beyond other approaches.

#### IMT (Identity Management Theory) and INT (Identity Negotiation Theory)

In contrast, Cupach and Imahori's IMT focuses chiefly on the role of identity, claiming that communication requires the participants to "successfully negotiate mutually acceptable identities in interaction" (Cupach & Imahori, 2005, p.196). For them, face and social identity are used to explore intercultural communication. Ting-Toomey refers to INT to explain how identity, both group-based and person-based, interplays in all intercultural exchanges and that 'in order to understand....you need to

uncover ways to affirm (your interlocutor's) positively desired personal identity' (Ting-Toomey, 2005, p. 217). The theory is based on the presupposition that all humans seek a positive identity and given that intercultural communication is a mutual activity, both or all of the parties involved are, to a varying degree, mindful of the identity negotiation process. Each person wants to present a positive personal identity and the success of the intercultural communication is reliant on the mutual support for and recognition of that negotiated identity. This theory, however, does not fully explain the added dimension of unequal language skills in the identity negotiation equation. Ting-Toomey states that a competent identity negotiator is 'able to use multiple cultural frames of reference to interpret a problematic, cultural collusion situation' (2005, p.230). In the context of this research, the students were not able to draw from multiple cultural frames of reference to negotiate their identity in the way that Ting-Toomey describes. Furthermore, neither of the theories mentioned above relates to an assessment or interview context: both relate to face and negotiation of positive identity as being key to successful intercultural communication, but this is only part of the story.

#### General applications of CAT

As shown, there are several theories touching on various aspects of intercultural communication but none offer such a wide-ranging scope as CAT for the purposes of this research. The applications of CAT have so far been much wider utilised in the fields of social science, criminology and medicine than in EFL research. In a study of inter and intra-generational communication perceptions, McCann and Giles (2007) used CAT to reveal age as a determinant in accommodation strategies in two different cultures, Thai and American. In both cases, young workers felt the older generation to be non-accommodative. In the above case, CAT is used to examine how people use language differently according to their generation and may communicate 'in ways that are biased in favour of their own age group' or accommodating their own age group. CAT has also been used to shed light on social status perceptions. Gregory and Webster (1996) analysed

interviews for the T.V chat show host, Larry King, in terms of perceived guest status by studying voice pattern convergence as an indicator of accommodation. Their findings reveal that people 'of high status and power' demand to be accommodated and that King himself accommodates to these high-level guests and is accommodated by lower status guests.

Watson and Gallois (1998) studied healthcare professionals in interviews and used CAT as a tool to establish that 'Health professionals in interpersonal interactions were perceived to pay more attention to relationship and emotional needs and to use more nurturant discourse management and emotional expression' (p.344) in order to facilitate effective communication between medical staff and patients. Myers and Giles (2008) applied CAT to police officer interactions with minor offenders for both minor and major traffic violations. The results showed that the officer was rated less favorably along both cognitive and affective dimensions when s/he non-accommodated rather than accommodated the offender. The application of CAT in all of these cases was to explore the dynamics of interrelationships in an interview situation in order to inform and generate best practice or understand more fully the non-verbal, non-linguistic information that is transmitted, consciously or unconsciously, and can have an impact on the outcome of the interview.

#### Application of CAT in EFL contexts

Similarly in the field of EFL, whilst less widely used, CAT has been applied to examine interview assessments. Richards and Malvern (2000) applied CAT in a case study to analyse the effects of vocabulary use by two teachers and the extent to which they adjusted their vocabulary to the proficiency level of the students in oral assessments. They found that the teachers did adjust their linguistic behavior to the students' proficiency. It is not clear from their research paper whether or not the interviews were scripted but given that they were described as 'free conversation' (p. 260) it is unlikely. This is a real dilemma for test writers. If the oral assessment is to be valid, it needs to reflect authentic communication but genuine

conversation is difficult to achieve given that this is an assessment context, there is a power differential between assessor and candidate and the topics of discussion are prescribed to elicit gradeable language items. In Richards's and Malvern's case study research, it was found that teachers tended to adjust their vocabulary not only to individuals but also to the perceived norm or average of the whole class. They conclude that oral assessment test validity depends upon appropriate accommodation.

Chen & Cegala (1994) used a case study approach and looked at topic management, comparing the accommodation strategies of two native speakers from the same country with pairs from mixed nationalities and nonnative language speakers. Here the application of CAT revealed that nonnative speakers tended to accommodate each other more, thereby offering support and striving to arrive at shared understanding. Native speakers tended to develop topics as an information exchange, based on shared meaning and understanding. This suggests that interaction is more effective when both parties are from the same sociocultural and linguistic background. Nishida (1999) refers to this shared meaning as 'cultural schemas' which are 'generalised collections of the knowledge that we store in memory through experiences in our own culture' (Gudykunst, 2005 p.404). Nishida argues that these schemas guide the way we behave and that when we move away from our own cultural environment we try to make sense through reference to our own native culture. In this way, we adapt and change our cultural schemas and each time that occurs, our brain registers the result enabling us to communicate cross-culturally more effectively (Gudykunst, 2005). The schemas that we possess inform our reaction to, and behaviour towards, other cultures. Accommodation theory provides a means of measuring that behaviour. It will be suggested that the extent to which convergence and divergence strategies are used can have an impact on grade outcome, and CAT is the tool that is employed to ascertain the degree and impact of accommodation.

#### 1.4 .Nonverbal interaction

Convergence and divergence, as presented in CAT, could also be applied to paralinguistic features of communication. 'Body language' is defined in the Concise Oxford dictionary (9<sup>th</sup> edition) as: 'the process of communicating through conscious or unconscious gestures and poses' (p.143). Feldman, Philpott and Custrini (1991) refer to body language as 'nonverbal behaviour', the perceived success in which equates to social competence. Implicit in this is the notion that if interactants do not share the same perceptions of body language, each is likely, consciously or unconsciously to feel or consider the other socially incompetent or at least different. Non-verbal behaviour has therefore a key role to play in intercultural interaction where gestures and poses may have very different meanings to the individuals involved (Feldman et al. 1991. p.330). According to Matsumoto (2001) , we are all born with the same universal expressions that we learn to modify, based on social circumstance, and we learn rules according to our culture about how to perceive and interpret those expressions: anger, disgust, fear, happiness, sadness and surprise. Research conducted on blind individuals confirms that blind people express these emotions using similar facial expressions to sighted people acquired without the visual stimulants of culture. If cultural expression and non-verbal behaviour are elaborations of these universal expressions that are learnt in later life, then it follows that they may also be unlearnt.

The perception of both facial expressions and gesture is culturally sensitive. There are many examples of gesture that exemplify this point, such as the circle formed by the forefinger and thumb can mean everything is 'A' Ok in the United States whereas in some Asian countries this gesture has sexual connotations. It is the perception of this gesture that differs. Similarly, direct eye-contact can convey honesty and candour to some where others would perceive this to be arrogant and rude (Valdes, 1998, p. 67-72). Non-verbal behavior is not limited to gesture, however.



### Multimodal interaction

More recently, Norris' notion, (2004) of '*multimodal interaction*', incorporating both verbal and non-verbal aspects of communication, shows the complexity of communication. Norris argues that the term 'non-verbal' implies a form of subjugation where the verbal is of primary importance however "modes like gesture, gaze, or posture can play a superordinate or an equal role to the mode of language in interaction, and therefore, these modes are not merely *embellishments* to language." (Norris, 2004, p.x)

She concurs that aspects of communication such as gesture are very important but then takes this a step further by also taking into account the 'material' world in which communication takes place, referring to this as a 'disembodied mode' (as opposed to the embodied modes of language, gesture and gaze) including such things as dress of interactants and the environmental context. Multimodal interaction for Norris is about the perceptions, thoughts and feelings that people are expressing which may or may not be the same as those that they are feeling such that:

there is a constant tension between what a person consciously does and what that person expresses. Interaction, then, is the exchange of communicated (expressed, perceived, and thereby interpreted) experience, thoughts and feelings of participants (Norris, 2004, p 4.).

This notion is echoed in CAT in the convergence and divergence aspect of interaction that is both a conscious and unconscious feature of communication.

So, if, as Norris states 'modes like gesture, gaze or posture' are integral to effective communication, what happens if these aspects are unavailable or denied to the communicants? Interestingly, following the SARS outbreak in Hong Kong, David Coniam (2005) conducted a case study of the impact of candidates wearing face masks on their oral test score. Examiners were also wearing face masks – this is significant as both parties in the

interaction were equally 'compromised'. Coniam found that wearing the masks had no impact on the final scores of the students and that both parties had employed measures to overcome the perceived and possibly real constraints that the face masks created. For example, some candidates simply talked louder or slower, others focused on maintaining continuous eye-contact throughout the interview trying to compensate for the 'disabling' effect of the masks. (Coniam, 2005) They were in effect using compensatory measures to get around the 'face-mask' disability because, for them, non-verbal facial gesture was embedded in their normal mode of communication: they compensated because they perceived they were disadvantaged by wearing the face mask.

In culturally-differing communities where the wearing of face masks (veils) is the norm, such as amongst Muslim women, compensatory measures might not be manifest since gesture and greater eye contact differ in terms of communicative significance. There is little evidence to support this statement but it is presumed that whilst for the mask wearers in Coniam's study this was a temporary condition, for Muslim women who veil, they will have done so probably from the age of puberty or even younger and therefore might not exhibit any compensatory measures to overcome what would be normal, daily wear, or the compensation may be automatic. Generally speaking, Middle Easterners (that is not to say that all Muslims hail from the Middle East or that all Middle Easterners are Muslims) retain 'a formality of manner, particularly in initial social relationships' (Parker, 1998, p. 95). Eye contact, particularly for women, would be reserved for close family members and would be seen as too familiar a gesture, particularly between members of opposite genders. Parker argues that where formality pervades social customs and daily routines, this has a de-personalising effect on interactions. The 'how' of interaction, therefore, is equally as important as the content that is communicated. It is important for Muslims, both men and women, to show respect and one way of doing this for women, is to avoid direct eye contact. To date there is, unfortunately, very little empirical data on the impact of veils in oral assessments and the

consequent effect this has on both interlocutor and assessor in a mixed-culture context.

### The balance of power in interviews

In any interview situation there is an unwritten 'asymmetry' of power-base largely due to the uneven ownership of knowledge (of questions or of environment for example) and this is certainly the case in an interview where speaking skills are being assessed, especially in the intercultural context. Jenkins and Parra (2003) found that test takers who were able to modify the power dynamic through appropriate non-verbal behaviour and thereby reduce the interview asymmetry achieved higher grades than those candidates who were not as able to manipulate the test situation in such a way. They cite research that supports the findings of the evaluators in the study, those being Anglo- Americans. They refer to eye-contact being linked to assertiveness (Romano and Bellack, 1980) and lack of it as a signal of non-immediacy and lack of involvement, whereas for the Chinese, assertiveness is negatively valued. Similarly, smiling, head nods and gestures led to higher evaluations by interviewers. With regards pitch, a flat or narrow pitch range is regarded as passive, cold and withdrawn, whereas greater pitch variety is regarded as more dynamic and extroverted; nasality has consistently been rated by North American English speakers as unattractive (Jenkins and Parra, 2003).

Jenkins and Parra (p.106 ) concluded that ' test takers are initially limited by the interviewer's right to nominate topics and control the discourse', but that those candidates deemed most successful were those who were able to reduce this asymmetry by employing non-verbal strategies, and not necessarily (by implication) those who were linguistically more able. They continue that they 'believe that it is essential to raise student consciousness about the vital role of non-verbal behaviour in their talk in interaction' (p. 107) . This consciousness raising should be a two-way process by means of which both interlocutor and assessor are enlightened. Intercultural communication stands to succeed or fail on this very point; consciously and

subconsciously very different culturally-embedded strategies interplay when cultures converge. These need to be explored in order to facilitate an understanding of the complexities of such exchanges.

Lantolf (2000) developed the notion of 'activity theory' based on Vygotskian sociocultural theory. The theory 'sees mental behaviour as action. It concerns all aspects of action; not just *what* the person is doing, but also *how* the person is interacting with objects and/or individuals in the social environment, *where* the person is acting, *when* the activity occurs, and *why* (Lantolf and Pavlenko, 1998).

This theory questions the root of assessment in that traditionally in speaking exams, it is the examinee's performance in isolation that is assessed. The tester is responsible for the questions, following the rubrics of the test as objectively as possible such that all oral tests administered on the same day (as in for IELTS) should follow the same administrative procedures.

Activity theory would also be concerned with looking at the joint interaction between the interlocutor and assessor with regards to socio-cultural norms; the focus should therefore not be on the individual's performance but more on the activity itself:

analysed in terms of the individual's motives and goals as well as the culture-based rules of the activity system that the individual is following (Luoma, 2004 p. 102) .

Luoma goes further to suggest that assessments themselves, when viewed according to activity theory, are an activity in which participants simulate real-life language use but since they know that it is not a real-life situation but a test, a further set of expectations and norms is involved emanating from the sociocultural perspectives of the communicants.

The traditional methods of oral assessment involving measurement by linguistic competencies alone are being challenged (Kirkpatrick, 2007, Wallace, 2002, Melchers and Shaw, 2003, Kramsch and Thorne, 2002).

EFL assessment can no longer simply be seen as candidate performance in terms of, for example, linguistic accuracy and fluency: the intercultural communicative competence of both participants – assessor and candidate as parties in the ‘activity’ must also be addressed.

### Intercultural communication

Aspects of intercultural communication may have a direct influence on the outcomes of oral testing. Here cultures meet face-to-face in a formal and potentially stressful setting. The assessment of speaking can be viewed as a dynamic performance where each party has a role to play. Oral language tests are qualitatively different from Listening, Reading & Writing where responses are in reaction to the test instrument not a human being. Oral tests are both personality-bound and culture-bound as the individual personalities and cultural backgrounds of both assessor and assessed have an impact on the communication that takes place. Oral communication, as has been seen, is multilayered and has many different aspects, including both physical and psychological (Luoma, 2004, McNamara, 2000, Underhill, 2003).

On a psychological level, oral assessment can be viewed as a performance. Participants or players may experience ‘stagefright’, that is a varying degree of anxiety associated with having to perform, and this can include both assessors and assessed (Jones, 2004).

Presented with a need to engage with a foreign language, a learner, or anyone who has some ability in that language, will make mistakes at a much higher rate than he or she would in any other area of study or professed knowledge. Such a person will also be very likely to suffer a drop in self-confidence that is not simply linguistic, for one is unable to be one’s ‘normal self’ and establish equal relationships; one’s self-concept may be undermined and one is subject to ‘negative evaluations’ (Tsui, 1996, p.155 ).

Horvitz et al (1986) observe that people may feel they are reduced to a childlike state when they have to use another language, an experience fraught with potential embarrassment. Confidence may falter and speakers may not be able to express themselves as succinctly as they would do in their own language, having to paraphrase or use over-simplistic words to express what they are trying to convey. Added to this is the possibility that the person may be asked to take on a different persona, with which they are not familiar either socially or culturally, such as in role-play questions. Generally speaking in most Western, individualist cultures, candidates are more likely to have experienced role-play as an integral feature of learning a language within the current trend of Communicative teaching practices. In other cultures, role-play could be perceived as too informal a mode of communication, even disrespectful. For example, a candidate might be asked if they play any sports. Of course if the candidate has the vocabulary and the topic is within their own realm of experience they will talk about sport whether or not they participate in it or have the slightest interest in it. They are therefore taking on another persona which could be construed as acting or indeed lying. This may make some candidates feel very uncomfortable but at the risk of simply answering 'no' they will feel obliged to follow through with an answer, truthful or not, in order to 'show off' their language skills. This feeling of discomfort may stem from both personal unease and cultural unease and may lead to limiting their own language learning and performance.

### Test anxiety

Ehrman, in a large empirical study in the United States, concluded that anxious learners 'tend to limit risk (and may limit learning too)' in order to save 'face' (Ehrman, 1996, p.96). Other research into East Asian students, for example, point to the fact that often learning styles in the students own cultural background determine the level of their active participation. Charlesworth (2008) found that Chinese and Indonesian students in U.S language classes tend towards a reflective learning style and listen to others rather than suggest their own opinions or speak out. Cheng (2000) dismisses

this and argues that this is not as a result of cultural background but more to do with familiarity with the teaching methodology and language proficiency. This is clearly true of oral assessments: where students feel uncomfortable, they may limit their own performance, take fewer risks and as such present too simplistic an answer to questions for fear of making linguistic mistakes. It could, of course, work the other way and students might rise to the challenge and take the opportunity to express as much language as they can, the stress being facilitating rather than debilitating.

As has been seen, test anxiety can be a result of various factors: worry about loss of 'face'; cultural or personal lack of knowledge and experience in the role-play situation; lack of understanding and use of the considered appropriate non-verbal behaviour and reluctance to 'perform', amongst others. In mainstream psychology the phenomenon of test anxiety is classified as having three distinct areas: trait, state and situation-specific.

Trait anxiety is part of one's character or an aspect of a more serious disorder. State occurs only under certain conditions, for example when one is asked to speak to an audience. Situation-specific arises only when certain factors are present; for instance, one might feel confident making a speech in English but when one has to make a speech in one's L2 she or he will feel much more threatened ( Jones, 2004, p.31)

Personality, environment and the extent to which the speaker has a fear of making mistakes have an impact, therefore, on the test taker. Add to this an intercultural dimension where perceptions of nonverbal communication and convergence and divergence strategies come into play and the level of test anxiety is liable to deepen. Situation-specific anxiety has particular relevance to Muslim women who wear the niqab (face veil). If students are asked to remove the veil for the oral test, test anxiety is likely to be greatly heightened where they will feel both personally and socially compromised:

with the veil their personal and cultural integrity remains intact, without it they may feel much more threatened.

### Cultural context

A further contributory factor which may lead to anxiety in oral assessment results is the cultural context in which the teaching, learning and assessment takes place. The cultural background of both assessor and candidate and the context in which the assessment occurs can have a major influence on the outcome. It may not be a level-playing field. They may have very different cultural backgrounds and the examiner will have a greater knowledge of the target language (and possibly target culture) than the candidate.

In some cultures the concept of “face”, for example is very important, both in front of your peers and in front of your teacher. Scollon and Scollon (1995, p.195) define face in communication as ‘the negotiated public image, mutually granted each other by participants in a communicative event’. It is this ‘negotiated public image’ that really comes into play in oral assessments and, I would argue, that is often far from ‘mutually granted’ but rather, mutually misunderstood and misperceived in certain cultural contexts. Muslim females, for instance, often feel very uncomfortable divulging personal information as is expected in the first part of the IELTS exam. As with most oral assessments, including PET and First Certificate (both part of the Cambridge University suite of English tests) part one of the test is designed (by Western examiners) to put candidates at their ease and allow them to communicate about areas with which they are familiar, such as their home background, their hobbies and interests. This may have the desired effects in like cultures, however, in Muslim culture, communication, especially with a stranger is noted for its formality. Having to discuss personal issues in this informal manner therefore might be assumed to add further stress where the candidate may fear loss of face, particularly where Muslim females are being interviewed by Western males. The examiner, albeit unwittingly, may be raising the affective filter in broaching personal topics; the candidate is compromising their cultural identity in trying to comply with the test requirements in order to pass the test. The candidate



may perceive the examiner's questions as intrusive; the examiner may perceive the candidate as formal and unable to converse in what s/he perceives as a conversational manner.

Jones (2004) cites the Japanese syndrome known as TKS (taijin kyofusho - fear of social relations) as a possible factor linking the cultural context with test anxiety; the delusional fear of embarrassing or harming others through one's inappropriate behaviour may be aroused in interactions with non-intimate neighbours, acquaintances, classmates, and co-workers.

Japanese candidates too may feel discomfort in discussing personal topics with a complete stranger in the context of an oral exam. Oral assessments, specifically high-stakes international exams such as IELTS and PET have been written from within a western cultural context with that particular set of beliefs and values. When these values and belief systems are then transposed on to other cultural value systems, misconceptions and misperceptions are perhaps more likely.

The concept of face is about identity respect and is tied to our own self-worth and to the value we attach to the worth of other peoples' opinions, (Ting-Toomey, 2005 p.73). In all cultures there is a sliding scale on which the concept of 'face' sits and beyond that there are the personal and contextual aspects that have a bearing on the perceived 'face' or loss of it. Hence, in the construction of international assessments, 'face' must be considered if we are to avoid intercultural misunderstandings.

### Sociocultural context

One aspect of intercultural communication, and more specifically oral assessment in that context, is the notion of individualism and collectivism and how that impacts on the participants. Generally speaking, western societies could be referred to as individualist cultures where the individual's goals are more important than the collective good of the family or community. In collectivist societies, the converse is true where individual needs are subordinate to those of the main group. According to Hofstede

(1984), collectivist countries include East Asian, Middle Eastern and some Latin and South American countries. With regards to language testing, following Hofstede's model, those belonging to individualist cultures may exhibit more self-confidence and be more comfortable with a one-on-one test situation whereas those from a collectivist society may prefer group assessments. Similarly in patriarchal societies, male candidates may also exhibit more self-confidence and suffer less test anxiety as a result of being the dominant members of their societies. In a patriarchal, collectivist society, it could be argued that a single woman, faced with a one-on-one speaking test with, for example an examiner from an individualist culture, (possibly male) could, potentially face extreme test anxiety that would negatively affect the outcome of her test. Add to that the possibility that the female is veiled and that her examiner is unable to use the convergence strategies that s/he would normally employ to encourage and support the student by using, amongst other strategies, increased eye contact and facial gestures. Even if s/he did use these strategies, the candidate would not be able to mirror these strategies and be seen to be doing so. This student might be seriously disadvantaged, however, it may have less to do with the perceived cultural background of the participants and more to do with the individual's personality and schema.

Bell (2006), talking about diversity in Hong Kong classrooms, argues that not only can students be categorised by race/ethnicity but also their sexual orientation, disability, and cultural background and that students may differ in motivation, learning style, learning strategies, aptitude, economic status and past learning experience. Gunzenhauser (1996, p.4) says, on the cognitive plane, students range in intellectual capacity, process capacity, and learning orientation. Along the emotional/behaviour plane, students could differ in the degree of maturity, conformity, self-esteem, and motivation.

Hofstede's model is based on the idea of culture being a measurable, static entity aligned to a particular nation and one which mediates all social interactions. Furthermore, Hofstede's observations were based on research conducted in the 1960s and 1970s and so are arguably irrelevant to today's

multicultural communities. Signorini, Wiesemes and Murphy (2009) challenge this model of culture, in particular in the current context of globalization, and argue that culture is dynamic and flexible. Their recommendation is to look instead at ‘micro-cultures’ and to appreciate that language and culture, whilst closely linked, are not necessarily the main indicators of cultural identity. Culture is dependent on much more than simply language or nationality and it is a dynamic phenomenon. This presents a problem for assessment writers of international English language tests trying to create a valid and reliable generic test that is suitable for all contexts around the globe. This study will try to address the issue of whether it is feasible to have such tests of English that have universal applicability.

### 1.5 Current research: IELTS and IELTS-style oral assessments

The three most widely known international EFL tests are TOEFL (Test of English as a Foreign Language), TOEIC (Test of English for International communication) and IELTS (International English Language Testing system). IELTS is part of the Cambridge University suite of exams of which more than eight million are taken every year, in over 150 countries. TOEFL and TOEIC, are administered in more than 6,000 institutions and agencies in 110 countries. They are American based tests and IELTS is jointly produced in Britain and Australia. All of these tests therefore originate from what Hofstede would consider to be individualist cultures and are written in that context for students around the world in both individualist and collectivist cultures. This may have an impact on how they are perceived by both assessors and candidates to a certain extent, but, as already stated, other factors such as personality and individual experiences will also have an impact. They all test the four language skills, in slightly different formats with TOEFL and TOEIC being largely administered online now, whereas IELTS is still largely paper-based.

A huge body of research has been funded by the IELTS joint-funded research program, sponsored by the British Council and IELTS Australia. Cambridge ESOL also supports the research through assistance to approved researchers. Rightly proud of their research tradition, IELTS is, according to their own declaration, 'one of the world's most researched English language tests, ensuring that IELTS continues to be the test that sets the standard through its high level of quality, validity, security and overall integrity'. (IELTS Research reports, 2006, vol. 6, p.v). Amongst other criteria concerning the logistics of conducting IELTS research (in collaboration with experts in applied linguistics and language testing), the self-funded research must have: relevance and benefit of outcomes to IELTS and potential to be published for both IELTS and an international audience. The research generated by the IELTS body is clearly commissioned to specifically improve and monitor the test primarily for its own ends; it is a commercial enterprise and therefore must remain competitive whilst protecting the reputation of each of the highly-respected bodies involved.

In the IELTS oral examination, examiners arrive at a grade according to criteria set out in band descriptors that focus on oral fluency, range of lexis, structural and grammatical accuracy. Similar descriptors are used in the TOEFL test. These descriptors are designed to be universal and all-embracing and only feature linguistic aspects of communication. However, when 'east meets west' and occidental culture meets oriental culture as in for example, a Japanese candidate and a European examiner, is the test robust enough to accommodate cultural differences and perceptions of the dynamics of communication (for both examiner and candidate) or do the tests implicitly disadvantage candidates from socio-cultural backgrounds which contrast sharply with western norms and values? If so, do we need to rethink the concept of universal descriptors and/or add a fifth dimension to the assessment process taking into account the cultural context, the perceived expectations of both parties and the nonverbal behaviour that takes place?

Brown (2007a) has led research specifically focusing on the IELTS oral interview regarding interviewer style and candidate performance in terms of the effect of interviewer behaviour on candidates. Results showed that “easier interviewers” tended to shift topic more frequently such that the interview appeared more structured whereas difficult interviewers engaged in more conversational techniques such as interruption and disagreement. Two types of interviewer were identified: ‘one who makes fewer allowances and provides less support and the other who uses simple language..... and who provides more support and feedback (Brown and Hill, 2007a, p.56).

The perception here is that interviewer behaviour, not only in terms of linguistic questioning styles but also in terms of ‘support’, has a direct impact on candidate performance. Brown and Hill consider the impact of interviewer turns, feedback and length of interview as measures of support but non-verbal behaviour/paralinguistic features of communication/body language and multimodal interaction also have a significant role to play. Added to this is the cultural dimension and the value systems underpinning both the interviewer’s and candidate’s perceived behaviour. In scripted interviews where interviewers follow a prescribed set of rubrics, such as in IELTS, it would be expected that ‘support would not be verbal as examiners must not deviate from the script (other than by using variations in stress and intonation patterns). The ‘support’ therefore would manifest itself as body language, gesture and stance and as we have already seen, all these features are culturally sensitive implying that well-intentioned support may for some be at once, misleading, confusing and even insulting. O’Sullivan and Lu (2002) found no significant difference in candidate language where the interlocutor script was deviated from, implying that interlocutor frames could be less prescriptive and more flexibility could be allowed without affecting the validity and reliability of the test.

Other IELTS commissioned research has targeted specific parts of the speaking exam and its effectiveness as a measure of linguistic ability

(Ingram and Bayliss. 2004, Paul, 2004, cited in IELTS Research reports vol.7, 2007). This research looked at the applicability of IELTS grade outcomes as indicators of predicted success in terms of 'language behaviour' in students' university courses following the achievement of the required IELTS band and their accuracy as indicators of 'language behaviour'. The results supported the case that IELTS results are a good indicator of predicted language behavior although there did not seem to be a link between success in coursework tasks and the IELTS score.

In similar research into specific sections of the test, Elder and Wigglesworth (2003, cited in IELTS Research reports vol.6, 2006) investigated planning time and its benefits for students in part 2. Their results showed that the various planning times of none, one minute and two, did not really affect grade outcome but they recommended that a minute of planning time be retained in part 2 'in the interests of fairness and to enhance the face validity of the test' (2006. p.13). Where the research has been directed at discourse specifically, it has been in relation to the IELTS banding criteria of linguistic competence; behavioural competence is not touched upon (Brown, 2003, Seedhouse and Egbert 2004, Read and Nation, 2002, cited in IELTS Research reports vol.6, 2006).

Research that has been conducted specifically within the context of this case study on oral assessment is thin on the ground. Lantegne's (2007) work on speech and behaviour perceptions between Arab learners and expatriate native English speaking teachers reveals how easily the line between behavioural norms and cross-cultural pragmatics can be confused. Her work has implications for teaching and testing because in the area of cultural norms related to language use, unstated (unacknowledged) differences in expectations can also lead to difficulty in communication. Lantegne (2006) also conducted research, specifically in the UAE, but as part of a larger project into semi-structured interviews with the premise that 'many English tests based on Western culture are inappropriate for regions where English use differs from that of Europe and North America'.

Her work focused on the wording of tasks rather than the language and behavioural interaction between the interviewer and test-taker, but nevertheless her conclusions do draw attention to the differences in English usage internationally and how that is not necessarily reflected in international EFL assessments. She maintains that the wording of a task affects its effectiveness. Lanteigne identified three tasks of non-Western language use that she felt to be culturally more appropriate in some regions than others. One of these included the topic of home food preparation. The writing task, although not specified in the research paper, relates to the fact that in several regions of the world domestic servants are employed to do the home food preparation. The underlying assumption was that it was normal practice to have house servants to do the cooking. The English teacher respondents, who were from 10 different countries and included both men and women, were almost equally divided as to the cultural appropriacy of this task. It is assumed that the task was considered 'culturally inappropriate' by 13 respondents because of its offensiveness and for being 'outside of (their) established socio-pragmatic use' (Lanteigne, 2006, p1). This highlights not only the fact that the wording of a task in terms of English usage, can be culturally inappropriate but also the fact that the task must have a socio-pragmatic meaning for the respondents: it must be possible to complete the task from within their schemata. Lanteigne does not specify whether the tasks were completed and so there is no evidence supplied to support or refute the theory that the output would be affected by the wording. However, Lanteigne's findings are supported by research into the importance of task authenticity. Ellis (1990 p.195) argued that linguistic competency is developed through meaning-focussed language activities that are appropriate to the learner's 'real operating conditions'. Long and Crookes (1992) also argued that linguistic tasks should relate to 'real world' tasks from within the learner's own scope of experience.

Khoddami's (2006) case study research, as an insider researcher, into English language teaching in Iran, uses Byram's (1997) model of Intercultural Communicative Competence (I.C.C) to highlight the role

cultural understanding has in ELT and how when learners are restricted from learning about the target culture, in tandem with the target language, they may be disadvantaged. This has implications for these learners when taking internationally recognized English tests and poses the question; does the responsibility lie with the Iranian teachers to explicitly teach 'culture' in the broadest sense or should international test creators consider adapting their tests to be less culture-specific, if they are found to be so?

Dahan (2007) discusses this idea in terms of whether teaching English means teaching its culture or including the culture of the learner. Her research, focusing on EFL in the Arabian Gulf region set out to test several hypotheses, one of which concerned students needing English mainly to communicate with other non native speakers of English and a second, that the students in the Arabian Gulf were not learning the language because of any interest in the target language culture. The findings were significant in that the majority of respondents do not need to use their English with native speakers and 92% of respondents had little interest in learning about the culture. Again, together with Lantaigne and Khoddami's Iranian perspective, the research suggests that the language that is taught and tested should take into account the context and the needs of the participants who will be using the language and therefore does not necessarily mean that the target language culture has any place in how the language itself is taught and tested. Furthermore, given that the students in Khoddami's study, will not be using their English with 'native speakers', to what extent does the oral English assessed need to match norms of native English speech and further, what are the 'norms of English speech'?

### World 'Englishes'

Recently, the concept of 'World Englishes' has emerged and the question that is being asked is whether or not there remains a model or yardstick against which spoken English can be measured. Without a model or point of reference, grading criteria such as those used by IELTS and ones created by colleges and schools around the world would lose credibility. The model or



norm to be used, however, is disputed. The International English (IE) view (Quirk, 1990) is that really only native English speakers represent the 'norm' whereas the World English (WE) view (Kachru 1992) is that this can prejudice and be discriminatory against non native English speakers. However, as Davies (2003) points out, there seems to be little consensus as to what each model constitutes but what is important is to identify those differences and to investigate how far the differences matter linguistically and how they affect understanding, intelligibility and attitudes. Furthermore, in terms of test scores, to what extent do these differences have an impact on grade outcome.

Lowenberg (2002) highlights the increasing problem of defining the use of English as a second language or a foreign language as in Kachru's (1990) defined 'Expanding Circle' and this has implications for the testing of English. Where English is used as 'foreign language', in countries such as Greece, Turkey, Japan for example, the norms of native speaker use from the inner circle would apply for testing purposes, Lowenberg suggests. In contrast, in countries where 'nativisation' (p.431) occurs and English is used by a large proportion of the population as a second, or even official language, the norms of language use may be different and so this too has implications for international language tests. Should test writers accept that there are different norms and models of English and adapt the grading criteria accordingly? Furthermore, as the demands for certification in English grow globally for employment, migration and further study purposes, so too does the need for accredited examiners to administer and assess the candidates.

It follows that a growing number of examiners will come from those countries where nativisation of English has occurred. Whether the test writers favour an IE approach over a WE approach, this has implications for test design, administration, assessment criteria and, not least examiner training. If a test is to be administered by an examiner from a country where nativisation has occurred and that test has been written to IE standards, this

adds another layer to the context, especially, but not exclusively, for the oral test. It effectively would be a test of IE, conducted (examiner) and received (candidate) by two participants for whom the language being tested is not necessarily the language use that they experience in their daily lives.

Lowenbourg refers to this as a 'blurring' ( p. 434) between the Inner, Outer and Expanding Circle defined parameters and suggests that tests developers need to be cognizant of this change in order to retain the validity and reliability of the tests they create.

Globalisation and the spread of English has been well documented (Pennycook, 2007, Fairclough 2006, Phillipson, 2008) and these studies reveal that there are ever-increasing numbers of people worldwide using English to communicate with other English speakers but also with other non-English speakers. Kirkpatrick (2007) , raises the issue of linguistic prejudice and this is of direct relevance to the case study. To what extent are test-takers subject to linguistic prejudice in the grading of spoken English, particularly in relation to the grade for pronunciation. Kirkpatrick links prejudice with social and cultural change; with the context. Whether we view English as Phillipson does as a 'cuckoo' or worse a 'lingua frankensteinia' (2008), the fact remains that in the Middle East region and in the U.A.E specifically, a good command of English, British or American, is considered a vital skill for employment and further education prospects. In China, British English has been usurped by American English as the 'prestige' accent (Kirkpatrick & Xu Xi 2002, cited in Kirkpatrick 2007) and this is due to socioeconomic factors related to changes in trading patterns. It is, in short, the *context* of the use of English that determines its variety. This case study focuses closely on the context of the study participants and its relevance in relation to their grade outcomes in IELTS-style oral assessments.

## Conclusions

This literature review presented the case that there are many different considerations that need to be explored further to try to address the main research question, namely: is it possible to have an EFL oral proficiency test that is universally appropriate and valid in all contexts? One of the starting points was to identify what is actually being assessed and therefore what constitutes 'communication', both verbal and non-verbal (Madsen, 1983, Fulcher 2003, Goffman, 81, Goodwin 1990, Philips 1972). Secondly, the cultural context of EFL teaching and testing, for both examiners and candidates, was explored (Kramsch, 1986, Halliday 1993). Following on from that, the methods, means and goals of assessment of that communication were discussed in relation to the IELTS test, current research undertaken in that field and the context of diversity and globalization (O'Loughlin, 2007, O'Sullivan 2000, Brown and Hill 2007a).

To sum up, there is no doubt that the demand for English language certification is growing and will continue to do so in the near future. As has been shown, there are concerns about what actually constitutes proficiency or competency in English and to what extent these are subject to the regional variations and circumstances in which, and purposes for which, English is being used (Kirkpatrick, 2007, Wallace 2002, Melchers and Shaw 2003).

These concerns directly affect how test writers decide on grading criteria and topic choice, for example. This case study will look at these issues specifically in relation to international oral assessment and further will consider not only the mechanics and administration of the test, but also the dynamics of the communication that takes place in order not simply to produce a speech act, but a co-constructed shared meaning between the examiner and candidate in the context of an intercultural exchange, via the medium of English. The case study will be in the context of a 'micro-culture' (Signorini, Wiesemes, Murphy. 2009), rather than applying Hofstede's (1984) more static model of a culture bound by nationality or

language, and will shed light on the many dynamic complexities that contribute to the act of communication.

The case study approach was chosen for those very reasons. The participants belong to a small contained 'community' within a single sex (students) college environment and have their own 'microculture' firstly as a whole group, then another micro-culture as Emirati females: 'inter-microculture' (their behavior towards each other) and an 'intra-microculture'. The microculture, as revealed in the oral assessment interviews is unique and as Yin (1994 p.13) states, a case study 'investigates a contemporary phenomena within its real life context'. Furthermore, Rowley (2002) argues that case studies are good for researching 'contemporary events when relevant behavior cannot be manipulated' (2002 p.17). The rationale for using a case study approach is therefore that the study is of a defined group of participants within a specific context responding to the challenges of an external input, that is the international English oral assessment. The case study, following Richards and Malvern's (2000) case study into accommodation in oral assessments, will focus on the following the key question: *International EFL oral assessments, does one size really fit all?*

In support of this, the following sub-questions will be addressed.

- 1. What role does 'accommodation' play in the negotiation of meaning in oral assessment?
- 2. To what extent can the IELTS publically-available marking criteria really reflect true communicative ability?
- 3. To what extent do the examiner's choice of topic and the student's 'world knowledge' help or hinder their performance in the test?
- 4. To what extent does the wearing of a veil impact the perceived quality of communication and the grade outcome?
- 5. Is examiner gender, in this context, a factor affecting candidate grade outcome?

## **CHAPTER 2: METHODOLOGY**

### **2.1 Case Study**

As has been seen, much of the IELTS -specific research thus far has tended to focus on the test, or the examiners or the test-takers in isolation and to a large extent the whole picture has not been taken into account. All of these aspects are undeniably important, but in the light of globalization and the international application of this test, it is now important to look more closely at the relationship between the test, the participants and the context and ask the question: Is it really possible to have an all-embracing test that is culture/context neutral and therefore of universal applicability? A case study approach was decided upon because of the context and nature of the research. The participant students are ostensibly from one cultural background in the sense that they are of the same nationality, gender and religious group but, as we have seen, culture is not so easily defined so it would be better to say they are from the same demographic group. In contrast, the examiner participants are from a variety of different countries, ethnicities and are of both gender. The setting for the research is therefore a micro-culture that has its own dynamics and was best suited to a case study approach.

Cohen et. al. (2000) remark on the strength of case study research as being able to establish cause and effect, and as being able to observe the context of those causes and effects. However, one drawback of the case study approach can be with regards to internal validity and reliability. According to Merriam (1998), reliability depends upon the extent to which the findings can be replicated and therefore triangulation of data is important. In this case, triangulation was achieved through observation schedules, questionnaire data and focus group data. Furthermore, feedback from participants was sought at various stages in the research cycle and colleagues were asked to collaborate and corroborate research findings in relation to the observation schedules. It was important to the researcher that all of the participants in the study felt that they had access to the findings

and were able to offer feedback and all were fully informed, from the outset, of the goals and parameters of the study.

The context of this case study, like those of Dahan (2007), Khoddami (2006) and Lanteigne (2006), is unique in that it presents an insight into a culture that is currently misrepresented and possibly misunderstood. The key protagonists are Emirati female Muslim students working towards IELTS exam success and expatriate English speaking examiners.

## 2.2 Socio-cultural context of the Case Study

The case study is set in the United Arab Emirates. It is an Islamic State run according to Sharia law, where equal proficiency in both Arabic and English is associated with success not only in the professional sphere, but also in the social and familial environment. Young graduates are encouraged to be bilingual both in their private lives as well as their public lives. It is commonly thought that only by being bilingual, English/Arabic, can Emirati nationals, men and women, take their place as modern citizens able to compete on the international stage. Due to the intense period of development of the U.A.E since the late 70s, there has been an influx of foreign workers of diverse ethnic and religious backgrounds resulting in the common language spoken between all the expatriate residents and the local population being English, not Arabic. The local population in the U.A.E, and in the Middle East region generally, is a young one. According to Lock (2008), 'people under 29 make up six out of ten of the region's population'. In the UAE, 51% of nationals are under the age of 19, 26% under the age of 9, (2010, UAE yearbook). In the UAE, as everywhere, the use of the internet as a means of communication has increased exponentially and so too the use of English as a result. English is used to communicate amongst Emiratis and non-Emiratis alike. And so are Arabic, Punjabi, Gujarati, Pashtun, Talagog and Hindi. The context for the case study is key to understanding the sociohistorical context of the participants, which is one of the main assumptions of CAT. Only by understanding the extraordinary

sociohistorical context of the participants can CAT be effectively used as an analytical tool.

The research project was conducted in two stages. Firstly a pilot project was conducted, on a smaller scale, to test the effectiveness of the research methodology and following a review of that, the main research project was conducted.

## 2.2 Pilot project

Prior to conducting the main research project, a pilot project was undertaken to assess the validity, scope and effectiveness of the research tools, practicalities of the data collection and the methods of analysis. The results from the pilot project shed light on the research process and helped hone the focus in the tools used in the main project. In the interests of research development, it is therefore useful here to first discuss the pilot project, the results of which helped inform and shape the main research project.

The starting hypothesis for the pilot project was that international English tests may not be objective reflections of communicative ability and that the administration, delivery and test constructs may disadvantage some candidates, in this case Emirati females. The fact that both participants, examiners and candidates come from contrasting cultural backgrounds may have a direct impact on the communication that takes place and the assessment process could be affected.

### Issues and concerns regarding EFL language tests

In most EFL oral assessments, the marking criteria cover the four key language competencies: lexis, cohesion, fluency and structural accuracy (see Appendix 1 for a copy of the publicly available IELTS oral band descriptors).

Within the Higher Colleges of Technology (HCT) college system, the context of the study, the grading covers:

- 1.communicative range - cohesion and range of speech
- 2.overall fluency – speed and pace
- 3.accuracy and appropriacy – grammar and vocabulary
- 4.pronunciation, intonation and stress – accent, L1 interference

The purpose of oral assessments is to assess the candidate's ability to communicate effectively. The performance during the test situation, it is hoped, will be an accurate representation of the level of communicative ability in a non-test situation. The criteria set in TOEFL, IELTS and HCT tests are intended for international use, for generic application regardless of the gender, religion, social status, cultural background of either the assessor or the candidate.

The aim of the pilot project was to explore the administration, delivery, test constructs and marking criteria of oral English tests to try to assess the existence of, and if so the level of, cultural bias and its effects on grade outcomes. It involved a series of 5 recorded interviews between Emirati female students (Bachelor of Education) and 4 English teachers, including both male and female, taped at the Sharjah Higher Colleges of Technology (HCT) women's campus. The interviews were based on the IELTS assessment and served as mock exams for the students. As such, the format and rubrics were familiar to the faculty and students. However, for grading purposes, HCT language banding scales were used as these were what the students were familiar with at the time. The interviews were conducted in a small, non-teaching room. In addition to the interviews, questionnaires were sent out to faculty and students asking them to consider the importance of knowing one's own and other cultures in the teaching and learning of language. There were also two focus groups, one involving the examiners and the other the candidates.

The main data collection tools used in the pilot project were, therefore, video recorded interviews, questionnaires and focus groups.



## Pilot project data collection

### Interviews

The interviews took place in a small room, with examiner and student facing each other across a hexagonal desk. The room was normally used as a storage area and so was not particularly comfortable or sound-proof, but was adequate. The materials used in the pilot project were all created specifically for this set of interviews and were therefore unfamiliar to either the assessor or the candidate, just as would be the case in a 'live' exam. As with IELTS, the assessor was asked to follow the rubrics of the assessment exactly and not to deviate from the scripted questions, as far as possible. Unlike the standard assessment, this interview was video recorded. All the interviewers in the project were qualified IELTS examiners and all the students had gone through several practice interviews of a similar kind in preparation for the actual IELTS exam in May. The B.Ed students must achieve an IELTS Band of 6 in order to graduate from the program. None of the interviewers had taught the students before and so the day of the interview was the first time they had met, just as it is in the 'real' exam.

There were four interviewers; one male and three female. All of the interviewers were European: one Irish, one Scottish and two English. They were all experienced teachers who had taught and tested within the college. All of the candidates were Emirati female, B.Ed students wearing the full length black abaya (cloak-like covering), two also wore a full facial veil. Prior to the interview (approximately 15 minutes) the interviewer was given the assessment pack to look through and select the questions that they would use in the interview.

The interview was both audio and video recorded. The candidates were graded as part of the mock assessment process, according to HCT oral banding criteria. Following the interview, the recordings were digitized, viewed and rated according to an observation schedule.

### Questionnaires

Questionnaires were distributed to both faculty and students via email. Some were returned as hard copies, others by email. Return was disappointingly low (56%) possibly due to the fact that distribution took place close to the end of the semester when exams were also taking place and both faculty and students were extremely busy. The questions were written with answers ranked (Likert method) to facilitate data analysis and to focus on strength of opinion or attitude to the statements on the questionnaire. The questionnaires contained two sections: one focusing on background information of the respondents, the second on perceptions of the value of knowledge of culture and cultural backgrounds in the teaching and learning process.

### Focus Groups

To ensure triangulation of data, focus groups were established of students and faculty to further explore the questions set out in the questionnaires. Two separate focus groups took place for students and one with English teaching faculty. The discussions were recorded and a summary of the discussions is included. The students who took part in the focus groups were not exclusively those who had been filmed, but all had responded to the questionnaires.

### Data Analysis

#### Interviews

For each interview the following six features of verbal and non-verbal communication were observed for both interviewer and student in each section of the interview; parts A, B and C.

1. Eye contact – regularity and length of contact
2. Hand movement – regularity and extent
3. Head movement – regularity and extent
4. Bodily stance – in relation to the other party
5. Clarity of voice – audibility and tone

## 6. Delivery of voice - speed, intonation, pronunciation

These features of communication were isolated to try to shed light on their impact on the communicated message. As Bull (2001) argues however, non-verbal communication is not always intentional and can even occur against the wishes of the speaker. He cites the example of trying to stifle a yawn when you want to appear interested. Non-verbal behavior (intentional or unintentional) and speech together deliver a message and this is referred to as 'mixed syntax' (Slama-Cazacu, 1976). Bull concludes that 'nonverbal behavior is so closely synchronized with speech that it should be regarded as part of natural language' (Bull, 2001, p.647). In the context of the case study, eye contact was identified because of the crucial role it plays in communication and several of the participants were veiled or partially veiled. Similarly, hand gestures were identified because of the role they play in Arabic in support of the spoken word.

Observation notes were tabulated for each interview and a colleague was in attendance for the video viewings to moderate these observations prior to the recording of written notes. CAT was the theory upon which the observations were based. Observations of these linguistic and paralinguistic features were observed in an attempt to measure the convergence/divergence strategies employed by both student and assessor.

### Discussion of results

Both the interview data and the questionnaire data presented more questions than were answered. In the interviews, two people from two different cultural backgrounds met and exchanged information in an intercultural interaction. However, this was an artificial interaction because it was a staged, mock exam. So, my first point is that whilst the intention was to study intercultural communication, the setting and context of the observation was already compromising the validity of the data.

Furthermore, before commenting on the various manifestations of 'cultural language' it is important to point out that in this context, both student and

assessor would be playing a role. The assessors were all trained IELTS examiners, which, at the very least, would entail them having to assume an official persona as a representative of IELTS. The students would be mindful that they were being assessed and videoed and that could affect their natural communication. Having said that, there were still some very interesting features that were common to all five interviews and would seem to indicate a trend or norm, even though five interviews is not enough to be able to make meaningful generalizations.

It was quite noticeable that all the assessors had a tendency to nod as an indication that they were listening to the student. It is difficult to assess whether this feature was a part of 'euroculture' or whether it was simply because the teacher was asking the questions and then had to seem to be listening, the nod confirming the latter. The students did not nod their heads as repeatedly as the teachers, and when they did, it was almost as a form of punctuation to their verbal utterances.

Secondly, it was very clear that hand gestures were central to the communication of the students, especially for the two veiled candidates for whom it could be argued that the hand gestures were compensating for the lack of visible facial expression. Interestingly, three of the assessors towards the end of the interviews were also beginning to use their hands more to supplement their utterances; indeed one teacher even mirrored the exact gestures of her student.

Eye contact was an issue for the male assessor with the students, not for them as might have been expected. Veiled students used direct eye contact with the male assessor, who presented as being very uncomfortable with this at first. As far as audibility was concerned, the veil had no detrimental effect on the clarity and volume of the students' voices – again, interestingly, it appeared that the male teacher when interviewing the veiled student, increased his voice level, possibly as a subconscious attempt to

breach the barrier presented by the veil, albeit a voile of paper-thin muslin cloth.

### Application of CAT

If the CAT model is applied to these interviews, where the teachers are expressing themselves more with hand gestures, it would appear that convergence is taking place. For the male teacher, however, there appears to be a clash of convergence and divergence: the students are using direct eye contact, unexpectedly, diverging away from what the assessor might expect but converging towards what the students feel is expected of them as English-language students. He appears to be diverging from his own cultural background and converging towards the students' cultural background. With regards physical proximity and stance, both veiled students sit upright and rather more formally further back in their seats whereas the non-veiled students tend to lean forward more to bridge the 'cultural gap' and converge with the perceived expectations of the target culture: their non-veiled status perhaps allowing them more flexibility and less constraint than their veiled counterparts.

### Marking criteria

The grades awarded to the students were the expected grades of between 6.5 and 7.5, the norm for this year group. In this set of interviews, there seems to be no pattern between lower/higher grades and gender of interviewer and whether or not the student was veiled, partially veiled or not veiled.

The Banding criteria used was HCT banding (see Appendix 2) but, similar to IELTS and TOEFL they are problematic to administer as they depend upon subjective judgements being made. For example, in Band 9, purportedly native speaker level, the Pronunciation criteria state 'only a slight accent may be noticeable': what accent, Northern English, Geordie, Indian and noticeable to whose ears? Sensitivity towards accents is a very subjective criterion. With regards the HCT Banding criteria in Bands 6-8, there are similar problems with interpretation.

In the 'Communicative Range' category, students are judged as to how well they can 'communicate on a wide range of academic and nonacademic topics which relate to own experience and interests'. In a relatively closed society such as is experienced by female Emirati students, their social experiences are limited compared with those of a European counterpart and thus their responses to questions about their free-time activities will have a narrower range. These students do not go to clubs, many have never visited a cinema and are always chaperoned if they visit shopping malls with other female friends. Therefore, the 'range' of non-academic experiences upon which they can draw is limited. Secondly, by whose standards are topics being judged academic and non-academic? Thirdly, how is the examiner to know the 'interests and experience' of the students, particularly a male examiner of a female candidate in this particular context.

In Band 9, of the 'Overall fluency' category, candidates are expected to be as fluent as an 'educated native speaker': an educated speaker, in whose terms and in which cultural context?

In Band 7 of the Accuracy and Appropriacy category candidates are allowed some 'lexical inappropriacies' – again by whose standards? In Indian English, often echoed in Emirati English due to the close historical trading links between the two countries, it is accepted that 'my leg is painning me' (my leg hurts) is perfectly accurate and appropriate English. It would be interesting to record how many British/American/Australian English examiners would perceive this to be a lexical inappropriacy or a grammatical inappropriacy. Another example would be the use of will (Irish English) and shall (British English) as in, 'shall/will I get you a cup of tea while I'm at the cafeteria?'

Finally in the 'Pronunciation, intonation and stress' category terms are quite difficult to quantify. For example, what is 'inappropriate intonation'? What does 'intonation and stress patterns approach native speaker level' mean?

Which standard are we measuring against? Pronunciation varies greatly between areas within Britain, let alone between America and the U.K. Clearer criteria need to be established.

Oral communication descriptors are very difficult to write but the evidence shows (see Appendix 2 for HCT Band Descriptors) that current assessment tools lack cultural sensitivity, an acknowledgement of 'World English' and are open to subjective interpretation. This applies not only to verbal communication but also at the level of non-verbal behaviour. The oral Band descriptors no reference to non-verbal communication strategies, the focus is exclusively on linguistic competencies.

#### Questionnaire data

##### Background information

The questionnaires reveal a difference in perception of the importance of culture between 'Western' faculty and Emirati students. It is important to note, however, that in some cases the students may have misread or misunderstood the questions as the questionnaires were not translated into Arabic and were completed entirely in English. The teachers, all English faculty working at the college, have collectively spent many years living and teaching in various parts of the world and have been exposed to a variety of cultural backgrounds. All speak at least one other language and all of them, both male and female, currently teach female students.

##### Importance for teachers/examiners of knowledge of the cultural background of students

The consensus of opinion was that it is important to know about the cultural background of the students (although what was actually meant by cultural background was not discussed) as a matter of respect, to be aware of behavioural expectations and to avoid misunderstandings. Most (six) thought that they knew about the cultural background of the students but for some this was limited to a generalized Arab culture rather than specifically Emirati and one teacher pointed out that cultural knowledge is often based

on hearsay particularly here in the Emirates, where there is little social mixing between Emiratis and non-Emiratis.

### Importance of the knowledge of the cultural background of teachers (for students)

As far as the students are concerned, they feel slightly less strongly that they need to know about their teachers' cultural backgrounds although they agree that doing so helps avoid misunderstandings. They also feel that they know the cultural backgrounds of their teachers; one student felt that she had learnt this knowledge from the movies! Another student remarked that her teachers do not mention their own backgrounds and so it is difficult to learn about them.

Teachers felt that it was important for students to know about their cultural background as they use study methods, assessments and resources based on a western culture and one teacher stated "mutual understanding is crucial in intercultural communication." Students also felt that it was important for teachers to understand their cultural background so that they could understand their responses to different situations.

### Knowledge of own culture

Both students and teachers felt they knew their own cultures well, despite many of the teachers having lived and worked abroad for several years. However, feelings were mixed as to whether we need to teach/learn about 'culture' – the students thought that it was important but the teachers had mixed views with one teacher pointing out that in English, there are very many different cultures that could be taught.

### Physical context for language learning

Regarding physical context for language learning and teaching, generally teachers and students felt that language must be taught within an appropriate context but not necessarily in the target-language country. Again, students and teachers agreed that it was important for the teachers to know a little of the students' language although the reasons given were



different. Teachers felt that knowing Arabic could be useful in picking up errors in translation and first language interference whereas the students felt that it would be useful for grammar explanations, encouraging shy students and for communicating with their parents.

### Teaching resources and materials

When asked about resources and materials the students felt that knowing the target culture was important but could not really quantify the reasons why or assess whether or not their texts met their needs as being culturally appropriate for them or presenting the target culture in an accessible format. The teachers, on the other hand, had mixed views. One teacher distinguished between linguistic competence and the need to be culturally aware for international communication purposes. Another expressed the opinion that the students need to be more tolerant and open-minded towards the target culture, whilst another mentioned that course books seem to reinforce the idea that western culture is superior. The dilemma appears to be that, as one teacher put it, books can create barriers but they also need to reflect and relate to the global community.

### Overview and commentary

Overall, there seems to be common agreement that if you come from a certain culture then you 'know' that culture, even if you have moved geographically or physically away from that culture and moved to live as part of a different one. This seems to imply that 'culture' is perceived as a fairly static phenomenon and that once you have experienced a particular culture then you 'know it'. But is this really the case? Were the respondents (teachers) simply referring to stereotypical perceptions of culture and would they now, as expats, still identify themselves with that culture or do they feel they have evolved, culture therefore being a dynamic phenomenon? Similarly, are the students confusing knowledge of their own traditions with knowledge of culture – after all the UAE has seen dramatic lifestyle changes in the past 30 years that must have influenced the culture; the values, beliefs and customs of the country.

## Focus Groups

### Students

The students' definitions of culture were heavily influenced by the importance of one aspect of their culture: Islam. Islam encompasses a code of conduct representing beliefs, values and customs and so is a driving force behind Emirati culture as a Muslim nation. Beyond that, symbols of their culture included camels and falcons, the National dress of black abaya and shayla for women, the white dishdasha and gutra for men; the eating style of eating on the floor using hands as utensils. In a similar vein, their interpretation of what represented British culture included clothing, food and celebration days such as Mother's Day. For neither culture was there a mention of language.

### Oral Assessment

On the topic of oral assessment, the students had surprising comments to make about their experiences. It might be assumed that female, Muslim students would be more uncomfortable in a one-on-one interview setting with male examiners from a different culture than with their fellow Muslim countrymen. It was unanimously agreed that an interview with a male Emirati examiner would be the worst possible scenario as far as feeling uncomfortable and intimidated is concerned. They felt the least uncomfortable scenario would be a female examiner closely followed by a Male, non-Muslim of non-Arab nationality. Part of the reason they cited was based on the religious belief and other, cultural norms. In Islam, it is forbidden for females to talk to males who are not family members or familiar to them and certainly not on a one-on-one basis, without the consent of their male guardians. The other feeling expressed was that Emirati and other male Arabs might want to dominate a discussion and take control, even stop an utterance to correct the perceived mistake. Furthermore, the students felt that this feeling of discomfort would be the case irrespective of them wearing the veil (hijab) or not.

In this case, therefore, it could be concluded that it is the cultural (pan-Arab) balance of power between the genders that the students feel is

affecting the nature of the interview and not, as might be presumed, the religious constraints.

### Faculty

The faculty definition of culture focused more specifically on value systems, a set of shared beliefs and a sense of 'grouphood', that is, belonging to a group of like-minded people for example, atheists or a political party. They all felt that culture is a dynamic phenomenon that moves with time and experience.

### Manifestations of British culture

A meal of 'Fish and chips' was one of the symbols representing British culture, according to the group but the question was posed 'what is British'? Many British passport holders are originally from different parts of the world and, may have different religious beliefs, social customs and behavioural norms. It was agreed then that culture does not correlate with nationality and that as different cultures sit alongside one another, new cultures may evolve such that 'culture' is a dynamic, moveable feast.

### Western culture in the Middle East

It was felt that Western styles, if not specifically British culture, were pervading every dimension of the media from advertising, T.V, printed newspapers and magazines. Furthermore, with regard to the college itself (modelled on a Canadian education system and staffed largely by expatriate faculty) it was felt that western influences in education were particularly noticeable. The example of the English tests used were cited as an example (PET, IELTS). In both tests, it was felt that symbols of western culture predominate and underline the message that to succeed in English language, knowledge of 'English' culture, as expressed in the Cambridge exams of PET and IELTS, is essential. In addition it was felt that this dominance of British/English culture in a country such as the UAE hosting many other variants of English such as Australian, American, South African and Indian was ill-founded.

### Emirati culture

A comment was made that it was ironic that Emirati nationals, albeit the ethnic minority in the population, were being pressurized into learning and even communicating in English in their own country. Emiratis are told that English is the international business language and that personal and professional success depends on their ability to communicate in English. Significantly, Arabic is not taught in the college, neither is Balooshi, a widely spoken Bedu language.

The faculty team expressed concern that this promotion of “west is best” (culture and language) was effectively eroding Emirati culture at an alarming rate.

### Overview and commentary

The focus group sessions were deliberately very loosely structured so as to allow participants to express their thoughts and reasoning about the issues. I felt it necessary to keep the focus centred on ‘culture’ but that it was important to allow the participants to take the concept and explore the issues that were pertinent to them. My role was as facilitator rather than interviewer. The focus groups sessions took place after most of the questionnaires had been collected in and it was clear that the usefulness of the data resulting from the feedback on the questionnaires was limited. The focus groups highlighted the need to be far more specific in terms of my research questions and suggested a reevaluation of the methodology for the next round of practical research data gathering.

### Review of pilot study methodology

The questionnaires did little more than provide a context and weak background insight into the students’ and teachers’ perceptions of what culture actually means and its relationship to language and language acquisition. This was because the questions were too broad in scope and did not hone in on the particular context of the case study. Also, it would have been more helpful to focus directly on oral tests themselves and query how the students and teachers understand what is meant by the marking criteria,

what do they understand by various levels of English ability, such as beginner and proficiency and what it means to them to be a 'native speaker', both linguistically and non-verbally.

Furthermore, the focus groups should have been far more tightly controlled so that a set of questions was drawn up first to guide the discussion, rather than letting it going off at a tangent. The dilemma here is that there is a danger of over-prescribing the flow of the discussion and therefore there is the risk of bias in that the researcher can lead the discussion in a direction that fits their needs in terms of data collection. The opposite is true too, that if the discussion is too broad in scope, the data required will not be generated although other data, previously not considered might come to light. A fine balance between guiding the conversation and allowing it to develop in different directions, not only between researcher and participants but also between participants was what was required. The purpose of the focus groups was to get a sense of the different perspectives on oral assessment that did not necessarily come though from the questionnaires.

Finally, while the interviews were clear and well-recorded, I think that there were too few to draw any conclusions from and that the scope needed to be much wider – more male interviewers (of various nationalities) with veiled, partially veiled and non-veiled candidates, as well as female interviewers. Also it would have been useful for the recordings to be viewed and graded by other assessors so that the grades could be compared. This would have added to the validity of the process and the comparison of grades may have shed light on the differences in interpretation of grading criteria and perception of how these are applied.

In short, the aims of the project needed to be clearer; the questionnaires did not reveal as much useful data as anticipated and the focus groups lacked direction; the interviews were well conducted but less well analysed. Whilst CAT was used as a tool for analysis, it was not clearly defined in terms of the specific attributes or expressions of language or behavior that would be

identified as indicators of the various CAT strategies; namely convergence, divergence and maintenance. There is a lack of research available demonstrating the practical application of CAT and this made it difficult to apply as there was no model to follow or adapt that was similar in scope to the case study context.

#### Implications for the main study

The pilot study was useful in that it was a chance to try out the methodology before the full scale research study took place. A number of issues arose.

- Questionnaires: it was debated whether to translate the questionnaires into Arabic but it was decided, for both the pilot and main study, that translating the document might cause confusion. The researcher, not being an Arabic speaker, would not be able to check if the translation matched the English version. Secondly, the pilot study questions did not focus directly on the topic choices available in the test and were generally too broad in scope. In the main study, a grid of topics was presented for respondents to indicate relevance or degree of difficulty that the topic might present. This was a major improvement as it focussed the data generated which was then directly relevant to the research questions.
- Focus groups: in the pilot study these were too loosely focused. In the main study therefore, the focus group questions were directly linked to the data that had been collected so that participants were able to offer their own feedback and input in to the analysis.
- Interviews: in the pilot study the recording process went well and this was repeated in the main study. One short fall was that only the researcher completed the observation sheet – in the main study, a sampling of the recordings were observed by a colleague who offered input and alternatives to the observations made by the researcher.

- Use of CAT: in the pilot study this was problematic for thereasons already cited. The observation schedule used was refined to include the questions asked and the observed accompanying non-verbal responses in an attempt to link the linguistic prompt with a non-verbal reaction as a means of evaluating the extent of accommodation taking place.

The case study approach proved to be the most appropriate method in the pilot study and was therefore followed in the main study.

## **2.3 Main Research project**

### **Methodology**

The research methodology was that of a Case Study involving a particular group of participants defined geographically, culturally, linguistically and by gender in a very specific context; a female campus of a technical college in the United Arab Emirates. The purpose of the methodology section of this thesis is to shed light on the *process* of the research.

The methodology section will present the following:

- 1. Conceptual framework for the research
- 2. Ethical considerations
- 3. Participants
- 4. Data collection tools
- 5. Analytical tool used: Communication Accommodation Theory

### **Conceptual framework**

The methodology of this research is largely interpretivist in nature in that it is looking at the way language and meaning are constructed whilst acknowledging that this can differ depending on the situation and the interaction between people and their interpretation of the situation. It is a qualitative study involving a case study including assessment interview data and questionnaire data. The case study here is socially constructed by me in the sense that I have gathered the participant volunteers and so probability sampling has been used. I needed to have some students who would normally veil for interviews and so whilst not identifying specific students, those chosen from the pool of volunteers included four who were veiled students. This, it was hoped, would enable me to draw generalizations applicable to the broader group of students. All the interviews were conducted in the same room, specifically set up for the filming of the interviews and so bias of location was avoided. The questionnaires were issued to all students in year 2 of the Higher Diploma program and therefore



were representative of the students in this college, in that particular year group.

The research design was largely inductive in that it was driven by data collection, data analysis and then subsequent development of theories based on the outcome of that data analysis. However, based on my own ontological practice and observations, it is also true that the research was slightly deductive in that I already had my own theories which I hoped the data analysis would be able to prove or disprove. I would suggest that the deductive and inductive models are not mutually exclusive and indeed the model shifted in emphasis as the research progressed. Wengraf (2001, p.3) refers to a “research cycle” in which researchers shift emphasis from deductive to inductive models at different moments in the research process. In the case study, the cycle began as deductive as a result of defining the research questions but then became inductive through the data collection process. The data analysis and discussion of results can be seen as both deductive and inductive as the conclusions drawn are as a result of the application of analytical review on empirical data.

The research was qualitative in nature, using the inductive grounded theory approach of systematic data collection and, in this case occurring within a small, contained, specified research context. The goal was that the theories arising as a result of the data analysis would shed light on current professional practice in EFL oral assessments and inform best practice both locally and internationally.

### Ethical concerns

Ethical research should consider the interests of the participants and all interaction with them should be respectful and fair and mindful of the context. Lankshear and Knobel (2004) list a set of ethical principles for good research: a valid research design, informed consent, avoidance of deception; minimised intrusion; confidentiality, minimised risk of harm,

respect, avoidance of coercion and finally, reciprocity. The college director was approached to ask for permission for the research to be conducted in the college. Not only was this willingly granted but he asked for a copy of the thesis to be held in the college library and in doing so, all the participants, including faculty and students and any other interested parties, will have open access to the findings.

The students all participated voluntarily and signed consent forms (see Appendix 3), countersigned by parents/guardians as would be legally required in the UAE for all females, stating that they were willing to be involved in the research and were fully aware that the interviews would be recorded. A guarantee was also issued that no part of any of the recorded data would be uploaded on to the internet and complete confidentiality of data was assured. Both faculty and students attended a brief meeting where the aims of the project were outlined and were given a chance to ask questions/opt out. The implications for those involved, especially in terms of time commitment, were made very clear. All interviews were to take place during the faculty and student 'free time'. For the students, they would gain extra practice in interview assessment with full and detailed feedback on their performance; the faculty would benefit from the discussion following on from the interview banding comparisons and of course would have access to all the speaking materials generated for the purposes of the research. All participants were therefore supporting me from a position of informed consent and were aware that their rights to withdraw at any time would not be challenged. The names of all the participants in this study have been changed to retain anonymity.

As an insider researcher, knowing, by face at least, all of the students involved and considering the faculty volunteers as good colleagues, it was important that I conducted the research in an objective and professional manner without jeopardizing the trust and respect of my students, colleagues or employer. At all times, the timetables of both student participants and examiners were strictly adhered to: if a student had agreed

to commit 20 minutes of her time, likewise an examiner, this time was tightly controlled in order to minimize intrusion on the normal daily schedule for both parties. In terms of reciprocity, the focus groups were designed to be 'loosely' structured to allow for free feedback on the research process and to facilitate input from the participants, both students and faculty, in the discussion of preliminary results. It was important that the participants had an opportunity to have an active input on the data so far generated, by them, and so whilst there were specific questions asked, there was also an opportunity for freer discussion. All participants were given the opportunity to comment on the final outcome.

### Participants in the study

#### Students

The case study was set in an all-female, all-Emirati, all Muslim (students) college of technology in Sharjah, an emirate in the United Arab Emirates. Sharjah is the "cultural capital" of the UAE and is known to be one of the most conservative. It is the only "dry" emirate, that is to say alcohol-free, emirate. The student participants in the study were all female Emirati, Arabic speaking students and the examiners were English speakers from various parts of the world.

For these students, IELTS exams are high-stakes, as already discussed, providing access to further and higher education (usually a Band 6 at IELTS is a prerequisite score for entrance on to a Masters programme in the U.A.E) or to their chosen career path. About 10% of the total student population attends college fully veiled with the face covered and therefore not visible. This percentage increases slightly in classes where there is a male teacher and within course programs that are seen as more conservative options, such as Health Science and Education. These are considered more conservative options as the job opportunities from these content programs are more likely to be within a female-only environment. Primary schools in the Emirates are segregated such that only female teachers would be

teaching female pupils and there are female-only departments in hospitals and laboratories.

The students involved in the interview stage of the research were all students in their second year of a Higher Diploma program, for them a Band 5.5 at IELTS is a graduation requirement. They sat the exam in May 2009 and as part of their preparation for the exam, would have had opportunities for practice interviews. An extra set of practice interviews took place in the last two weeks of February and first week of March for those students who had volunteered to be involved in the research project. All of the students in the year group were contacted by email to ask them if they would be interested in participating in the research project and being filmed whilst undergoing a full IELTS-style oral assessment interview. The students were therefore not pre-selected. Out of the whole year group of 176 students, 11 students finally gave their consent. Whilst this was not a huge number to be able to draw generalizations from, it was nevertheless a manageable size given the time constraints and demands on other people's time that this project commanded.

It was important to me that some of the students would normally veil in an interview but not wanting to bias the initial request for volunteers, this was not mentioned in the initial email. Three of the 11 volunteers wore veils during the filmed interviews, and at approximately 27%, this number represents a true picture of the % of veiled students (partial and fully-veiled) generally. Of those three, only one was completely veiled (whole face covered) and at 9% this is probably more than the average if the whole student population was measured. All students were at roughly the same stage in terms of English level but within the year group there is usually a broad range from weak Band 5s to strong Band 7 candidates. It was not my intention to deliberately filter ability ranges as ability did not have any bearing on my research – how the assessor arrived at the Band Level did, however.

### Faculty/examiners

In a report commissioned by IELTS in to the validity of the rating scales (bands) and the rating process, Brown (2007) found that examiners found the scales easy to interpret and apply, following an update of the format of the test in 2001. Brown highlighted the facts that different examiners may privilege, or give greater attention to different aspects of performance in oral assessments despite the four rating categories being equally weighted; that novice examiners may rate differently from experienced examiners and that examiners may deal differently with problematic interviews. It was also noted that some examiners may adhere strictly to the rating criteria whereas others may be more inference-orientated and respond to candidate strategies of humour and creativity (Meiron, cited in Brown, 2007).

With regard to the experience of the examiners, the training for IELTS examiners is strictly monitored so as to minimize the difference between novices and experienced examiners. Examiners are recredited on a two-yearly basis and must be examining regularly within that time to retain their examiner status. The areas that are not so easily controlled are that of the influence of inference and how closely examiners refer to the rating criteria to the exclusion of any other features, such as body language. One examiner may feel that grammatical accuracy is the key to successful communication and that eye contact is irrelevant (eye contact is not mentioned in the bands) whereas another examiner may feel that eye contact is crucial to the fluency of the exchange, for example, and grade negatively (consciously or unconsciously) if the candidate does not make eye contact.

The faculty examiners who took part in the conducting the interviews and in second marking the audio recordings were all faculty members from the women's campus and all were accredited IELTS examiners. Access to faculty members at the adjacent men's campus was possible but in the interests of trying to keep the variables as low as possible only faculty from the women's campus were approached. This was also a practical move as it was easier to accommodate requests for changes in interview slots or other

changes to plans from within the same campus. It also meant that all the faculty involved were familiar with teaching female students and were known to the students, although it was a deliberate policy not to have students assessed by an examiner who had also taught them English. This is in line with IELTS policy where examiners should not know the candidates they are assessing, ideally.

All English teachers who were also accredited IELTS examiners (19 in the college) were emailed and informed of the project, its scope and research questions and invited to volunteer. Their commitment was initially for one 15 minute interview and to listen to and grade an audio recording of a different interview; a further 15 minutes of their time. The same faculty were also invited to participate in the 20 minute focus group meeting. Nine teachers agreed to participate in the research. This meant that two of the interviewers would have to conduct two interviews, if all 11 students were to be interviewed. Once volunteers had been identified, a quick meeting was arranged for a question and answer session so that the logistics of the interviews could be finalized, according to faculty and student timetables. This was also an opportunity to thank faculty for their contribution, and reassure them that they could withdraw their support at any time and that no timetabled classes would be affected by the interview schedule. As mentioned above, all interviews were to take place during the faculty and student 'free time'. For the students, the benefits were that they had access to extra interview practice with experienced examiners who would give them immediate grade feedback. For the teachers, the benefits were that they would have access to free use of all the purposely created materials for the project and would feel empowered to offer their input in a research project the results of which could impact on their teaching strategies in the future.

## Data collection tools

Data triangulation was achieved by using three methods of data collection: mock interviews (audio & visual recordings), questionnaires and focus group meetings.

## Interviews

The interviews were conducted in an unused classroom, this facilitated being able to set up the cameras and recording equipment and leave them in situ. It also meant that the participants had one venue that was secure and there was no risk of this room being double-booked for teaching purposes or, once the “filming” sign was posted on the door, of being interrupted. A desk and two chairs, for the one-to-one interviews, were moved into the room and screens were used to demarcate the filming area, create an interview setting and maximize the light conditions. The audio recordings were recorded on digital recorders by the examiners. The same type of recorder was used as is typically used in IELTS exams so the examiners were familiar with the equipment. Two digital cameras were used, set at slightly different angles on tripods. This was done to ensure that even if the participants moved out of camera shot for one camera, they would still be filmed on the other. It also meant that the risk of losing the filming due to technical hitches such as low battery power or lack of memory space/film available was halved. The cameras were set up and filming initiated by the researcher, who then left the room once filming started. This was to reduce the possible impact of having a third party in the interview (which would not happen in normal circumstances) and to try to allow the participants to ignore the cameras and feel as relaxed as possible. The researcher remained outside the room, returning to switch off the cameras once the interview was over, as signalled by the examiner. All 11 interviews were conducted in this manner.

The interview materials used were created and compiled for examiners specifically for the research project (see Appendix 4). They were based on the IELTS format and resources were collected and adapted from a range of commercially available IELTS practice materials. The materials included a file of scripted questions organized according to the three parts of the speaking exam. For each part, choice of question topics was guided by what was available on the IELTS official website and published IELTS practice materials. It was very important that topics used were as close as possible to what might be used in an authentic exam as cultural bias in topics was one of the research questions being considered. It was also important that there was a choice of topics available, as in the live IELTS tests, as whilst this opened up the chance that the examiners might all choose the same topics, it also rendered the assessment more authentic. Had examiners all chosen the same topics, test security may have been compromised in the sense that candidates may have discussed the topics chosen in their interviews, thereby arming the next candidate with useful knowledge about what was coming up in the assessment. As the candidates came from a variety of content programs and had interviews scheduled at different times and on different dates, it was unlikely that this kind of 'leakage' of exam questions would have happened.

It was also very important that the rubrics and format of the test were as close as possible to that of the exam for the students so as to offer valid practice. The delivery of the materials during the interview was kept as close as possible to the real exam being mindful not to infringe IELTS copyright rules. Included in the interview pack was a set of publicly available IELTS band descriptors (see Appendix 1), a grading sheet, plain paper (for student use in part two of the interview), a feedback sheet for students to take away with them and writing materials.



### Interview data

All 11 mock interviews, set up to mimic the real conditions of an IELTS oral test, were video and audio recorded; each was transcribed (see Appendix 5). In the videoed transcriptions, notes were taken of the linguistic interchange but also of the physical, paralinguistic interaction taking place. For each interview a grade was recorded for the live interview by the examiner and then another was recorded by a different examiner based on the audio recording alone. In live IELTS exams, the interview is audio-recorded and the examiner at the time registers the grade on a grade sheet. The audio-recording is only revisited for auditing purposes to monitor examiner performance or to verify a grade where a candidate has queried the grade given. For the purposes of the research it was important to have the two grades to be able to assess whether the dynamics of a face to face interview where the examiner was actively engaged in the conversation, compared to the perspective of an examiner listening to an audio recording would have any bearing on the grade given. If there were grade differences (either positive or negative) between the audio recording and the live interview, this could help shed light on the degree to which non-linguistic communication adds to the comprehension and fluency of the communication taking place. CAT was the tool used to analyse and interpret this non-linguistic aspect of communication.

An examiner allocating a grade from an audio recording is limited to the information on the recording; the linguistic production of the candidate and any pauses, silences, umm-ing and ah-ing, laughter, sighing and coughing. The non-linguistic communication, sometimes referred to as paralinguistic communication, body language or gesture was observed and recorded in tandem with the transcription of the linguistic exchange (see appendix 5). The purpose of doing this was to try to align 'events' or expressions of paralinguistic communication with the context or linguistic expression of that communication. For instance, if a question posed appeared difficult for the candidate to answer, was it the linguistic (lexis, grammar, conceptual) content of the question or was it the manner in which it was delivered, both

linguistically and paralinguistically? Further to this, if the student was veiled and unable to express facial gestures openly, would this enhance or hinder the communication exchange enough to make a grade outcome difference, positively or negatively?

The observation schedules (see appendix 5) of the video footage were completed alongside the transcripts of the spoken exchange. The specific features of non-linguistic behavior identified included eye movement, head movement, facial gestures, posture (whole body), proxemics (distance from the desk/other person), hand movements, general fidgeting and 'other', such as coughing or clearing the throat. These universal aspects of non-linguistic behavior were then analysed using the parameters of CAT: convergence, divergence and maintenance to assess the impact of the non-linguistic behavior on the shared understanding that took place and the resultant grade outcome. The observation schedule was based on the model used for the pilot project but was adapted and extended in the light of the pilot project experience. The non-linguistic events highlighted in the observation schedule were selected by the researcher. It could be argued that this has the potential for bias, however, as a member of the same sociocultural background (microculture) as the students and examiners, the researcher was in a position to identify and therefore record the relevant non-linguistic behavior. Gallois et al (2005), in presenting CAT as a tool, point out that the researcher is indeed expected to identify the markers for accommodation and non-accommodation, but no guidance is offered as to how.

The grade outcomes were another measure used in the data analysis. The grades were awarded according to the publically available IELTS bands (see Appendix 1). Each examiner (face to face and audio) awarded their grade independently, without discussion with the second examiner. In fact neither examiner knew who was grading which interview. This added to the validity of the grading process. Given that each examiner was an accredited IELTS examiner, it was less likely that any grade difference would be due

to examiner error, examiner variation or misinterpretation of the grading schema and more likely that it would be a direct result of the two different contexts for the grading process.

The students in the study should all have been within the following IELTS

Band range: [http://www.ielts.org/PDF/UOBDS\\_SpeakingFinal.pdf](http://www.ielts.org/PDF/UOBDS_SpeakingFinal.pdf)

Band	Fluency & coherence	Lexical resource	Grammatical range & accuracy	Pronunciation
7	<ul style="list-style-type: none"> <li>Speaks at length without noticeable effort or loss of coherence</li> <li>May demonstrate language-related hesitation at times, or some repetition and/or self-correction</li> <li>Uses a range of connectives and discourse markers with some flexibility</li> </ul>	<ul style="list-style-type: none"> <li>Uses vocabulary resource flexibly to discuss a variety of topics</li> <li>Uses some less common and idiomatic vocabulary and shows some awareness of style and collocation, with some inappropriate choices</li> <li>Uses paraphrase effectively</li> </ul>	<ul style="list-style-type: none"> <li>Uses a range of complex structures with some flexibility</li> <li>Frequently produces error-free sentences, though some grammatical mistakes persist</li> </ul>	<ul style="list-style-type: none"> <li>Shows all the positive features of Band 6 and some, but not all, of the positive features of Band 8</li> </ul>
6	<ul style="list-style-type: none"> <li>Is willing to speak at length, though may lose coherence at times due to occasional repetition, self-correction or hesitation</li> <li>Uses a range of connectives and discourse marks but not always appropriately</li> </ul>	<ul style="list-style-type: none"> <li>Has a wide enough vocabulary to discuss topics at length and make meaning clear in spite of inappropriacies</li> <li>Generally paraphrases successfully</li> </ul>	<ul style="list-style-type: none"> <li>Uses a mix of simple and complex structures, but with limited flexibility</li> <li>May make frequent mistakes with complex structures, though these rarely cause comprehension problems</li> </ul>	<ul style="list-style-type: none"> <li>Uses a range of pronunciation features with mixed control</li> <li>Shows some effective use of features but this is not sustained</li> <li>Can generally be understood throughout, though mispronunciation of individual words or sounds reduces clarity at times</li> </ul>
5	<ul style="list-style-type: none"> <li>Usually maintains flow of speech but uses repetition, self-correction and/or slow speech to keep going</li> <li>May over-use certain connectives and discourse markers</li> <li>Produces simple speech fluently, but more complex communication causes fluency problems</li> </ul>	<ul style="list-style-type: none"> <li>Manages to talk about familiar and unfamiliar topics but uses vocabulary with limited flexibility</li> <li>Attempts to use paraphrase but with mixed success</li> </ul>	<ul style="list-style-type: none"> <li>Produces basic sentence forms with reasonable accuracy</li> <li>Uses a limited range of more complex structures, but these usually contain errors and may cause some comprehension problems</li> </ul>	<ul style="list-style-type: none"> <li>Shows all the positive features of Band 4 and some, but not all, of the positive features of Band 6</li> </ul>

To arrive at an overall grade for linguistic communication, a band is given for each of four categories: Fluency and coherence, Lexical resource, Grammatical Range and Accuracy and Pronunciation. The grades are then averaged. For example: FC 6, LR 6, GR 5, P 5, would generate a final grade outcome of **5.5**.

The application of these band descriptors to spoken output can be problematic for examiners for two reasons. Firstly, examiners are juggling with conducting the interview and following the time constraints and rubrics closely to maintain the validity of the assessment internationally, but at the same time they need to be engaging in conversation and listening carefully to the responses of the candidate. Furthermore, the descriptors themselves are quite difficult to apply to a speech act because of the existence of qualifiers such as “usually” and “generally” (see Band 5, FC or Band 6, LR). How often does something have to happen for it to be considered usual or a general pattern? Similarly, in the FC column, the criteria states “speaks.....without noticeable effort”. How can we be sure that every examiner is gauging “effort” in the same way, indeed is it possible to gauge effort? Is that effort associated with a physical ability or a mental ability? In the LR column, band 5, the criteria mention “familiar” and “unfamiliar” topics but we are not sure to whom this familiarity applies, presumably the candidate, but how is the examiner expected to know what a familiar topic is? The GR column mentions using “complex structures with some (or limited) flexibility”. Again, the term “complex structures” is contested and what exactly is meant by ‘flexibility’? It could, for example, mean that one or two structures are used correctly in a variety of sentences or that the candidate has been flexible in the broad sense of trying out complex structures but perhaps not successfully. It is very difficult to write descriptors that will cover all eventualities with clarity and precision, but where there are areas that require interpretation from the examiner; this is where the disparities in grade outcomes may be likely to occur.

## Questionnaires

Self-completion questionnaires (Byram 2001) were used whereby the students and faculty answered questions by completing the questionnaires themselves. The questionnaires were administered in the period after the interviews had taken place. Two questionnaires were created and distributed; one for faculty and one for students (see Appendices 6 and 7). Burgess, Siemenski and Arthur (2006) outline the advantages and disadvantages of using self-administered questionnaires: they are cheap, quick and easy to administer, however the data from them requires a lot of time to collate, present and analyse. Burgess et al point to the fact that one disadvantage is that respondents can read the whole questionnaire before providing their answers. This was not an issue for these EFL students as they needed to take each question step by step. Their language level dictated this and the researcher led them through the questionnaire slowly and methodically to ensure that all respondents understood the questions being asked and the range of possible responses in the multiple choice sections. The teachers (faculty) responded to their questionnaires in their own time and so it was difficult to monitor how methodical their approach was, or whether or not they discussed their answers with any other teachers. For the faculty questionnaires, the same team of 19 EFL teacher/examiners who were approached to be involved in the filmed interviews were emailed the questionnaire and sent a hard copy. Confidentiality and anonymity were assured. 10 completed questionnaires were received.

The questionnaires for both teachers and student were divided into 3 main sections: background information, language teaching context, varieties of English. It was important for the contextual understanding of the interview data to learn more about the attitudes, perceptions and opinions of the examiners via the questionnaires. It was hoped that the questionnaire data would shed light on the reasons for the choice of topics used in the interview and their perception of the cultural context in which they were teaching/examining.

The student questionnaire was administered slightly differently but similarly divided into the same three sections. To ensure like delivery, the researcher administered the questionnaires to all of the students involved. The questionnaire was written in English and not translated into Arabic. Teachers of all the sections (class or group) in year 2 were sent an email requesting 20 minutes of their class time for the questionnaire procedure. For various reasons, including timetable clashes, not all teachers were willing or able to do this but for those who could, the researcher visited the classes and explained the purpose of the research to the students and requested their support. It was made clear that contributions would be anonymous, confidential and were entirely voluntary. 74 completed questionnaires were collected. It was decided to administer the questionnaires in this way to facilitate collection but also to ensure continuity of delivery. The researcher could ensure that the same level of support could be offered to all students; that being linguistic help with lexis or question comprehension, without influencing the student responses and thereby ensuring data reliability. I had to be very cognizant of the fact that any responses to questions that the students might ask could have been coloured by my desire to retrieve particular aspects of data. In the pilot project, I had asked teachers to give out and collect the questionnaires and it became apparent from feedback that different teachers had taken very different approaches. I needed to be sure that each student answered the questions honestly and without discussion with their peers. On the one hand discussion can often be thought-provoking and allow for a variety of perspectives; on the other hand, students can also simply write down what their friend thinks and this has the opposite effect. The data from both sets of questionnaires were compiled and presented in pie chart and bar chart format.

### Focus groups

The third tool used for data collection was Focus group meetings. Two different sets of meetings were convened; one for the faculty and one for the students. In each case, the same sets of target groups were emailed (all of

the teacher/examiners in year 2 and all of the students in year 2). In the teacher focus group, there were teacher/examiners who had not been part of the filmed interviews, present and in the student groups there were students who thus far had not been involved. In both sets, therefore, there was a mix of existing research participants and 'newcomers'. This was welcomed as all perceptions, attitudes and experiences were considered valid. The participants were briefed on the initial data retrieved from the questionnaires; it was summarized and comments were invited and then a guided discussion followed.

As Kreuger (1994) argues, focus groups should be planned with a defined topic, in this case the interview and questionnaire data. The interviewer's role, who Kreuger refers to as the 'moderator, is to nurture the conversation flow. It was important that the participants had an opportunity to have an active input on the data so far generated, by them, and so whilst there were specific questions asked, there was also an opportunity for freer discussion. The focus group data were transcribed (see Appendices eight and nine) and are summarised in the Data Analysis section.

#### Analytical tool used: CAT

CAT is one of the many theories of communication that can be used to try to explain the dynamics of communication. Amongst those, there are three key theories that deal with accommodation or adaptation: CAT, Intercultural Adaptation Theory (Ellingsworth 1983) and Co-cultural Theory (Orbe, 1998). In this case, CAT was identified as the most appropriate theory as it broadly embraces the intercultural aspect of communication that exists in most international EFL oral assessments, the focus of this current research. It concerns itself with the shifts between interactants as they converge or diverge during their interaction and it is suggested that mutually convergent behaviour is observed when interactants are supportive of each other and that this facilitates communication. In EFL oral assessments, the hypothesis is that to maximize the possible grade

outcome for the candidate, mutual convergent behavior needs to take place because accommodation facilitates communication. The theory provided no framework for its application so this was constructed by the researcher. The design of the observation schedule (see appendix 5) was problematic because of this, but it was decided to record any significant non-verbal behavior (or non-behaviour) alongside a commentary of the questions being asked. The scope of the observations was deliberately broad in order to capture as much data as possible. It must be noted that these observations were based on the 'micro-culture' of the researcher themselves (and a colleague who confirmed or disputed those observations, as a secondary observer): as an experienced IELTS examiner, as an experienced English teacher working in the socio-cultural context and as coming from a Western background.

The responses to the questions were not recorded here as the main focus was to try to shed light on the candidate's behavioural responses to specific questions. The oral responses were recorded on audio files.

#### 2.4 Communication Accommodation Theory (CAT)

The data collected was analysed using CAT as a theoretical model and concerned three principal areas: linguistic communication; paralinguistic communication; the test itself (rubrics, context, topic choices, grade outcomes). A cornerstone of this theory is that speakers consciously or sub-consciously adapt the manner of their speech to convey attitudes, values and opinions and adjust their behavior accordingly. So, communication is seen as not simply linguistic and gestural expression but also behavioural expression. Convergence, divergence and maintenance are the three main strategies explored by CAT. The three strategies are further divided in terms of production, magnitude and reception. When a speaker attempts to *converge* towards the speech and non-speech patterns it is thought that the speaker is seeking the recipient's approval or is aiming for a high level of communicational efficiency. In the same situation, a speaker using a divergent strategy would be expressing a desire to communicate a 'contrastive self-image' or even to change the recipient's manner of speech



to converge more with theirs. The *magnitude* of the convergence or divergence would be measured by the speaker's ability to adapt and the level to which the speaker needs social approval and/or high communicational efficiency. For the recipient, *convergence* is positively evaluated when the speech style of the interlocutor is perceived to be close to their own or the perceived intent is altruistic or non-threatening and this leads to higher ratings for friendliness and even attractiveness. The opposite is true for divergence in reception which is negatively rated and can be perceived as selfish or distant. (Gallois et al, 2005).

#### Assumptions of CAT applied to the case study

I now return to the three CAT assumptions detailed on page 16, that communication takes place in sociohistorical context, is about both referential meaning and negotiation of identities and is achieved via accommodation through a variety of moves.

The first assumption underpinning CAT is that communicative interactions are embedded in socio-historical context, so it is important to establish what this is in terms of the case study participants. The student participants come from a homogenous ethnic background in that they are all Emirati females. That being said, many Emirati families have a history of intermarriage not only between Bedu tribes but also with trading partners, notably from the Indian subcontinent and Iran. Many Arab speaking Emirati families also include an Indian or Iranian female second wife so there is a strong linguistic and socio-cultural influence from outside the Emirati geographical borders with several different languages often being spoken in the home apart from Gulf Arabic; Farsi, Balochi, Hindi, Tagalog. The current Emirati population is therefore a hybrid of various ethnic groups very strongly linked to the economic and trading history of the region, via the sea and across the desert, such as in pearls, spices, building materials and cloth. The Emirati population, however, is a minority in its own country with current estimates of less than one third of the inhabitants being Emirati passport holders. The rest of the inhabitants have come from all over the

globe to live and work, each bringing with them their own language, culture and identity in to what is commonly referred to as the “melting pot” of Emirati society. The socio-historical context of the students in the study therefore is not quite so clear cut as it may seem at face value and it has changed with the rapid development of the country. The socio-historical context of their forefathers of pearl divers, nomadic tradesmen and fishermen has now been superseded by modernity and an oil-industry driven society where wealth and luxury are prized. Furthermore, against the background of a multi-ethnic expatriate workforce there is a strong drive to reinforce ‘Emiratiness’ and so the state-funded college, for example, requires its students to wear the national dress of abaya and shayla (for females) and is only accessible to Emirati students.

In contrast, the examiner/faculty come from a much wider variety of socio-historical contexts which are largely secular, and the countries for which they are passport holders are much older and established: Australia, Ireland, Scotland, England, the United States of America. This juxtaposition of the young fledgling socio-cultural context of the students set against that of the more established, mature socio-cultural context of the examiners is possibly one area that could impact on the convergence/divergence strategies of the speakers in an interview situation.

The second assumption is that communication is about both exchanges of referential meaning and negotiation of personal and social identities. In EFL oral assessment contexts, referential meaning may be less effective as referential meaning is derived from shared experience and knowledge – reference to something that is not within your experience or knowledge is less likely to result in effective communication. Similarly, if personal and social identities are very different, this too may have an impact on the effectiveness of the interaction. In the context of the case study, there is shared referential meaning in terms of living in the UAE but that is qualified by the status of being either an Emirati national or an expatriate. Emiratis enjoy a variety of benefits compared with expatriates in areas such

as property ownership, employment rights and so on and therefore are perceived (by themselves and others) as a privileged section of society. Expatriates, depending on their nationality, have different social identities. For example, to generalize, most expatriates of Asian descent are employed in the construction industry or in manual labour positions whereas expatriates from English-speaking countries tend to be employed in white-collar, higher status positions. Social identity is directly linked to the socioeconomic status, the gender, the religious and cultural beliefs and constructs of the individual. In the context of the Case study, the Emirati female students therefore have a very different personal and social identity to that of their expatriate examiners. The students are Muslim, none of the examiners are; the students are Emirati nationals, none of the examiners are; the students are of Arab descent, none of the examiners are; the students are females, some of the examiners are male. Furthermore there is an age difference: the students' ages range from 17 to 22 whereas the examiners are all 30+.

The third assumption underpinning CAT is that the extent to which interactants achieve the informational and relational functions of communication is determined by their ability to accommodate their communicative behaviour, through i) linguistic ii) paralinguistic iii) discursive and iv) non-linguistic moves, to their interlocutor's perceived individual and group characteristics. This is the main concern of this study, that is, the degree to which the interactants in an EFL oral assessment accommodate each other and the resultant effect that this may have on the grade outcome.

To sum up, a case study methodology was chosen as the most appropriate method of investigation for the context of this 'microculture' and CAT, although problematic to apply and operationalise, was nevertheless considered a suitable means of theorizing and analyzing the data that was collected.

## **CHAPTER 3: DATA ANALYSIS**

### **Introduction**

The analysis and interpretation of the data is set firmly within the context of the case study; a college of higher education, female campus, in the United Arab Emirates. The results will be discussed for each of the three data collection tools in the context of the key research question: does one size fit all when it comes to EFL oral assessment? It is hoped that the data will reveal evidence to illuminate the process that occurs in cross-cultural exchange, not only in linguistic terms but also behaviourally and how that impacts on the grade outcome of EFL oral assessments.

### **3.1 Linguistic Communication: Grading criteria**

The interviews were conducted as mock interviews but using similar rubrics and format to IELTS interviews and at the end of each interview, the examiners gave the candidates feedback and a grade according to the publicly available IELTS band criteria. The grades are shown below. All of the students were preparing for the IELTS exam in which they needed to achieve an overall Band of 5.5 (this average also takes into account their Reading, Writing & Listening bands) to proceed to their Higher Diploma course; 6 to continue on the Bachelor program.

Interview Number	Examiner F/M	Candidate V/PV/NV V=veiled PV =partially veiled NV = non-veiled	Grade (IELTS band) ¼ band college adaptation				
			FC	LR	GR	PR	Final
1 *Muna	Tony M	V- only eyes showing	6	6	6	6	6
2 Maysa	Caron F	NV	6	6	6	7	6.25
3 Sawsan	Caron F	NV	6	6	6	7	6.25
4 Saffiya	Neil M	V completely: eyes hidden	6	6	6	6	6
5 Shamsa	Bob M	NV	7	6	6	7	6.5
6 Faiza	Mark M	V- only eyes showing	6	6	6	6	6

7 Hanan	Ida F	NV	6	6	6	7	6.25
8 Amina	Hilda F	NV	6	6	6	6	6
9 Abeer	Tony M	NV	6	6	6	7	6.25
10 Sameera	Adrian M	NV	6	6	6	7	6.25
11 Fayrouz	Mary F	NV	6	6	5	7	6

\*All the names listed above are pseudonyms to preserve the anonymity of the participants.

As can be seen, all of the candidates are within the expected range for their year group and ability.

The linguistic communication of the candidates is evaluated using the banding criteria (see Appendix 1) and cover four key areas: Fluency & coherence; Lexical resource; Grammatical range and accuracy and Pronunciation. The communication of the examiners is strictly controlled by the rubrics of the test itself. IELTS is a scripted test with some flexibility to deviate from the script only in the third part of the test. All accredited IELTS examiners go through thorough training to qualify as examiners and grades are moderated to avoid inconsistencies where possible. The examiners are trained to adhere strictly to the rubrics of the test, and to the banding criteria.

In the Case Study, despite all the examiners being accredited, experienced examiners, there are clear differences between the grade outcomes of the audio and the video, or face-to-face recordings. If the rating criteria, as interpreted by the examiners, were entirely reliable, this would not be the case. Brown's study highlights the decisions that examiners have to make in interpreting the rating criteria and that those are influenced by the behaviour and attitudes of the examiners themselves. The extent to which this occurs can be explored, in part, by applying the principles of CAT. This will be discussed further in the analysis of the video footage.

### Audio grading

The same grading procedure was applied for the audio recordings in that the public Bands were used and the examiner recorded a score at the end of the interview. In each case, two different examiners graded the candidates; one grade was given for the audio only interview and the other grade was awarded by the examiner conducting the face to face interview. The results of the grades can be seen in the table on the following page.

In IELTS usually only whole or half grades are possible whereas the college, for coursework allows quarter bands for coursework purposes.

Face to face grades are shown in Black; Audio grades are shown in *italics*; the difference is shown in **BOLD**

Interview Number (in order of video transcripts)	Examiner F/M	Candidate V/PV/NV	Grade (IELTS band) ¼ band college adaptation					
			FC	LR	GR	PR	Final	Difference
1 Muna	Tony M	V: eyes showing	6	6	6	6	6	
	Mary F		6	6	6	6	6	0 M/F
2 Maysa	Caron F	NV	6	6	6	7	6.25	
	Tony M		6	6	6	7	6.25	0 F/M
3 Sawsan	Caron F	NV	6	6	6	7	6.25	
	Tony M		6	6	6	7	6.25	0 F/M
4 Saffiya	Neil M	V : eyes hidden	6	6	6	6	6	
	Ida F		5	6	5	6	5.5	- 0.5 M/F
5 Shamsa	Bob M	NV	7	6	6	7	6.5	
	Adrian M		6	6	6	7	6.25	- 0.25 M/M
6 Faiza	Mark M	V : eyes showing	6	6	6	6	6	
	Hilda F		5	6	5	6	5.5	- 0.5 M/F
7 Hanan	Ida F	NV	6	6	6	7	6.25	
	Caron F		6	6	5	6	5.75	- 0.5 F/F
8 Amina	Hilda F	NV	6	6	6	6	6	
	Neil M		6	7	6	7	6.5	+ 0.5 F/M
9 Abeer	Tony M	NV	6	6	6	7	6.25	
	Mary F		6	6	6	6	6	- 0.25 M/F
10 Sameera	Adrian M	NV	6	6	6	7	6.25	
	Mark M		6	6	6	6	6	- 0.25 M/M
11 Fayrouz	Mary F	NV	6	6	5	7	6	
	Bob M		6	6	5	6	5.75	- 0.25 F/M

Seven out of the 11 candidates achieved lower grades when assessed solely on the audio recording. Three candidates were awarded the same grades and only one candidate achieved a better grade. The examiners were all the same examiners who had conducted the face to face interviews; all were experienced and certified IELTS examiners so the grading process itself should have been uniform, or as uniform as would normally be expected. There are often discrepancies between examiners and it could be argued that a grade allocation by one individual alone is not a valid assessment. In research, triangulation of methods and approaches adds validity to the final analysis so perhaps there is a case for more than one examiner or method of assessment.

The differences in the grades, where at 0.25, are minimal. However in four cases, where the difference brought the grade below six, this was enough to fail the students on their oral performance at college level.

With regard to the gender of the examiners, as can be seen from the table, the discrepancies between the grades were varied:

Male Interviewer/Female Audio examiner (4): 0, -0.5, -0.5, -0.25 (average difference: -0.03125)

Female Interviewer/Male Audio examiner (3): 0, + 0.5, -0.25 (average difference: +0.083)

Male Interviewer/Male Audio examiner (3): 0, -0.25, -0.25 (average difference is: -0.16)

Female Interviewer/Female Interviewer (1): - 0.5 (average difference is - 0.5)

Firstly, it must be noted that any analysis of the gender here is not particularly representative given that there is only a small and not an equal number of examples for each category and no control group. Having said that, it is interesting to note that, for the figures available, it appears that the least average discrepancy between examiners occurs when there is a male interviewer with a female audio examiner (-0.03125). The greatest average grade difference appears between two female examiners (-0.5). The second highest average grade difference is between two male examiners (-0.16). This would seem to suggest, on first analysis, that in terms of grade



difference, it might be good practice to have both male and female examiners grading candidates.

Following on from that, if we consider the gender of the examiners grading in the face to face interviews, we can see that in the four interviews (Interviews 1, 4, 6 & 9) where males interviewed and females graded the audio recordings, three of the four male examiners gave higher grades and the fourth was an identical grade. In the reverse scenario, there is no real pattern; in interviews 2 & 3 there was no difference between the grades but in interview 8, the female face to face examiner graded the candidate lower and in interview 11, the interviewer graded the candidate slightly higher than the male audio examiner. In the three interviews where there was no gender difference between each examiner (interviews five, seven and ten), all of the face to face examiners recorded a higher grade than the audio examiners of between 0.25 and 0.5 of a band. Overall, in 10 of the 11 interviews, the grades for the audio recordings were either the same or lower than those for the face to face interviews.

Given that in this study, the same set of examiners was responsible for both types of grading, it can be assumed that the banding criteria were being applied consistently. In the audio recordings, only the linguistic data is being processed by the examiners – that is all they have to assess the candidates with and that is what the bands are designed to assess, linguistic competency. In the face to face interviews, I would argue, both linguistic and communicative competency is being assessed and this is not being adequately addressed in the band descriptors, hence the grade discrepancies between face to face and audio grading. If we now take a closer look at those areas where grading discrepancies occur (see table below), we can see that the number of discrepancies varies according to the particular descriptors. Grade variations are highlighted in shaded areas.

Interview Number (in order of video transcripts)	Examiner F/M	Candidate V/PV/NV	Grade (IELTS band) ¼ band college adaptation					
			FC	LR	GR	PR	Final	Difference
4 Saffiya	Neil M	V : eyes hidden	6	6	6	6	6	
	Ida F		5	6	5	6	5.5	- 0.5 M/F
5 Shamsa	Bob M	NV	7	6	6	7	6.5	
	Adrian M		6	6	6	7	6.25	- 0.25 M/M
6 Faiza	Mark M	V : eyes showing	6	6	6	6	6	
	Hilda F		5	6	5	6	5.5	- 0.5 M/F
7 Hanan	Ida F	NV	6	6	6	7	6.25	
	Caron F		6	6	5	6	5.75	- 0.5 F/F
8 Amina	Hilda F	NV	6	6	6	6	6	
	Neil M		6	7	6	7	6.5	+ 0.5 F/M
9 Abeer	Tony M	NV	6	6	6	7	6.25	
	Mary F		6	6	6	6	6	- 0.25 M/F
10 Sameera	Adrian M	NV	6	6	6	7	6.25	
	Mark M		6	6	6	6	6	- 0.25 M/M
11 Fayrouz	Mary F	NV	6	6	5	7	6	
	Bob M		6	6	5	6	5.75	- 0.25 F/M

### Lexical Resource

As can be seen, LR (lexical resource) has the lowest number of grade discrepancies at one, compared with three each in FC (Fluency and Coherence) and GR (Grammar range & accuracy). The greatest number of discrepancies occurs in the P category (Pronunciation) with five instances. To a certain extent, these results are not at all surprising. In the LR category, for example, the criteria are more tangible; the examiner is asked to check for the range and choice of vocabulary and note instances of, and the extent to which, less common and idiomatic items of vocabulary are used. These are countable features of linguistic ability. There remain grey

areas in the descriptors of course, such as “manages to talk about familiar and unfamiliar topics” (Band 5, public version). It is left to the examiner’s discretion as to what constitutes a ‘familiar’ or ‘unfamiliar’ topic. This is clearly problematic: how does an examiner know what the constructs of the candidate are? Even if there were common ground between them in terms of nationality, culture, life experience and age, this would be difficult, but where the two interactants are from a different sociocultural background and a different generation and possibly gender, the examiner has to rely on his or her own experience as a teacher/examiner. They draw on their own experience of what they think an EFL speaker should be able to speak about. Generic topics such as family, friends, work and social life are common in EFL materials for all levels from beginners through to advanced but what is a familiar topic to the candidate in their own language may become unfamiliar ground in the target language being tested.

### Fluency and coherence

With regards FC, band discrepancies occurred in interviews 4, 5 & 6. All three of the face to face examiners graded this category a full band higher than the audio examiner. In all three, the face to face examiner was a male and in two cases the candidates were veiled, one fully and one with her eyes visible. In interviews four & six, the audio examiners were female, in interview five, male. It might be suggested that in the cases of the veiled candidates, the veils obstructed the audio recording and distorted the sound quality. Having checked this myself, with another colleague, we found this not to be the case. None of the audio examiners knew of the veiled status of the candidates. As mentioned before, if we consider the band descriptors, examiners are asked to make quite subjective judgement calls with some of the criteria. For example, it is difficult to see how it is possible to ascertain what is “language-related hesitation” (Band 7), particularly if you are unable to see the person’s facial and bodily gestures that might lend a clue as to whether the candidate is mentally searching for “content” or “words or grammar” (Band 9). Here the descriptors make a distinction between whether fluency is affected by the candidate’s ability to talk about the topic

area, the “content”, or familiarity with the language tools of words and grammar. I would suggest that the two are very often interlinked and again very difficult for the examiner to judge either way. Further, the descriptors in band 9 also talk about developing “the topics fully and appropriately”. This is another example of subjective criteria – what constitutes “fully” and “appropriately”? “Fully” could mean discussing a part of a topic in minute detail or the whole topic in very broad terms; interpretations of ‘appropriately’ could vary greatly depending on the sociocultural context of the exchange and the individuals involved.

### Grammar range and accuracy

All three of the differences in grades between GR occurred as higher grades for the face to face interview than the audio grading. Interestingly, all three ranges of veiling were evident here: interview four (veiled, eyes hidden); interview six (veiled eyes visible); interview seven (non-veiled) and with two different interviewer gender ratios: four & six M/F; seven F/F. This would suggest that GR is probably the least affected of the grading categories in terms of gender and veiling but, surprisingly, all three were a band lower when audio graded. As with LR, it might be assumed that GR is more finite and tangible, countable even and so there would be no added communicative value assignable to a face to face interview. With this limited data of three interviews, it is impossible to make generalizations but this is an area that would benefit from further research. In the GR category, examiners are asked to assess whether the candidate “frequently produces error-free sentences” (Band 7) or “may make frequent mistakes with complex structures”. There is still the issue of the use of qualifying adverbs (frequent and usually, for example) and subjectivity but mistakes and error-free sentences are quantifiable. It is therefore surprising that there would be consistent difference between the two grades. When assessing the audio recordings, the examiners are able to focus purely on the linguistic data and apply the grading criteria without having to enter into a conversation with the candidate. Perhaps when the examiner is involved in the conversation, they are less likely to focus their attention towards grammatical error and

more consider the global communicative exchange grading the communication on the effectiveness of the grammar in terms of conveying a message rather than on the countable linguistic features uttered.

### Pronunciation

The grades for P had the highest number of discrepancies at five. All of the candidates were non-veiled and four of the five grades were higher for the face to face interviews than the audio versions. Pronunciation is arguably the hardest category to grade, not least because the public band descriptors are incomplete; full descriptors are only available for Bands 2, 4, 6, 8 & 9. Secondly, pronunciation, more than any other category, is subject to the prejudice of the examiner in terms of what is considered 'correct' pronunciation. Fulcher (2003) makes a distinction between pronunciation of single words and at phrase or sentence level arguing that at single word level, mispronunciation is distracting but rarely leads to miscommunication. He suggests that the purpose of the pronunciation style is more important, for example, a newsreader may be expected to pronounce a particular word in a certain way, but he challenges the idea that pronunciation need even be assessed in EFL tests at a single word level.

For Fulcher, the choices speakers make in terms of tone and intonation are more likely to affect communication. Luoma (2004) points to the fact that most learners are assessed for pronunciation against a perceived native-speaker standard and asks if that standard is justified. The point is made that most learners would fail to achieve native-like status even if they are fully functional in the target language in terms of communicative ability. Luoma calls for assessment of pronunciation to take account of this and be 'guided by native speaker standards but defined in terms of realistic learner achievement'. (2004, p.10)

Examiners may have a preference for American intonation patterns, for example, or British English vowel articulation and this may in turn influence their grading. In Band 8, the descriptors read 'is easy to

understand throughout; L1 accent has minimal effect on intelligibility' (Appendix 1). This is extremely difficult to quantify; examiners who have spent a lot of time teaching a particular nationality, for example, Japanese students, may have much less difficulty in filtering through the L1 accent than those who have not. Mispronunciations are perceived amongst native speakers and discussed at length; take the word 'scone' for example, or the word 'route' (Davies 2003, Gass and Lefkowitz, 1995). In interviews seven to eleven, apart from eight, the face to face examiners graded higher than the audio examiners and this would seem to confirm that other features such as non-verbal communication acts, in addition to the bare linguistic articulation of words and sound, play an integral part in the interaction.

### Zero grade difference

In interviews one, two & three there was no grade difference between the audio and face to face grading despite the differences in interviewer gender and student veiling status. In interview one, the examiner was a male and the candidate was fully veiled with only her eyes showing. In interviews two and three, the same female examiner interviewed both non-veiled candidates. In these three cases, it could be argued that the grading rubric is fully applicable and valid for both audio and face to face grading. However, it should be noted that these three interviews only represent 27% of the total findings and, whilst significant, are not representative of the general trend.

### Overview

The examiners for both the audio and face to face grading were from the same pool of experienced, certified IELTS examiners so the application of the grading rubric should have been consistent. If the grade differences are calculated as a whole, 73% of the grading was misaligned in some way; 9% higher for the audio grading and 64% lower. When IELTS conduct moderation of interviewers, a single band difference in one category is considered acceptable, so, in all of these cases if the interviews were 'live' the grade differences would have been acceptable variances. It is interesting to note that all moderation of live IELTS interviews is conducted via an

audio recording. If candidates wish to query the score they receive for the oral assessment, the recording of the interview is what has to be used for the reassessment, unless the candidate wishes to go through the whole process again.

These findings, although conducted on a very small scale are nevertheless significant; 73%, almost three-quarters, of the grading was different despite the examiners being from the same pool. If the same methodology was applied on a much wider scale and other variables were factored in such as, for example, experienced against newly qualified examiners, native speaking examiners against non-native speaking examiners, the time of day of the assessment and examiners from different cultural backgrounds, the percentage difference in grades could potentially be even higher.

Significantly, the area with the most frequent grade discrepancy was in Pronunciation with five instances of grade differences. Critics of the veil might assume that it would be veiled candidates who would score lower in P due to obstruction of the mouth or distortion of sound as a result of wearing the veil; this was not the case. In both of the interviews of veiled students, the grades for face to face and audio recordings were the same. It would seem more likely, based on the evidence shown above, that grade differences in all of the categories are a result of the difference between the examiner perception of the interaction occurring in face to face interviews and of the evidence available on the audio recordings. In the audio recordings, the examiner is merely an observer. The face to face interview is not only a dynamic exchange of speech but also of cultures, past experiences, expectations and aspirations as communicated through body language, facial expression and physical interaction with the environment and each other. The question therefore needs to be asked: is it realistic to be using the same grading criteria in both scenarios? If not, can the grading criteria or grading process be adapted accordingly? Should there be a separate set of criteria for audio recordings?

### 3.2 Face to face grading and the application of CAT

To reiterate, CAT is based on three general assumptions: that communicative interactions are embedded in sociohistorical context, that communication is about both exchanges of referential meaning and negotiation of personal and social identities and, that interactants achieve the informational and relational functions of communication by accommodating their communicative behavior through linguistic, paralinguistic, discursive, and non-linguistic moves, to their interlocutor's perceived individual and group characteristics. (Gallois et al, 2005, p. 137)

In analyzing the interaction during the interview, several variables were observed and transcriptions were made for each interview. Apart from the linguistic exchange, the following features were noted: eye movement, head movement, facial expression, posture, proxemics, hand movement, general fidgeting and 'other'. If convergent accommodation occurs, the supposition is that this would be observable in terms of mirrored or like behavior, through facial expression or other non-verbal behavior or in the manner of speech, for example, pausing, altered speech speed, intonation and rephrasing. The video footage has been transcribed so that observable, physical events can be cross-referenced against the spoken interaction and then this in turn, with the grades awarded, veiled status of the candidates and the gender of the examiners.

When applying the assumptions of CAT, we can assume that the female Emirati candidates have a relatively culturally uniform sociohistory in that they are all Muslim, Emirati, Arabic mother-tongue nationals who have lived in the Emirates all of their lives. However, as we have already seen, the situation may be more complex: culture is not necessarily defined by nationality or language. Their examiners are non-Muslim, of non Emirati origin and non-Arabic speakers who may have lived and worked in a variety of different countries. The two parties are therefore quite distinct and have quite different social and personal identities. Similarly, as discussed in the literature review, both parties have different 'perceived individual and group characteristics'. The Emirati students are familiar with



their matriarchal society that has an emphasis on group support and joint responsibility; the examiners are from societies that value individualism and sole responsibility.

In three of the interviews there were no grade differences between the audio grades and the face to face interviews so it could be assumed that the non-verbal communication that took place was of little or of no consequence. On the other hand, it could be that both parties accommodated to each other equally and thus there was no imbalance and this is manifested in the zero grade difference. In the first interview, the examiner was a male with a veiled candidate (eyes showing) whereas the second two were conducted by the same female examiner and neither candidate was veiled. Below is a tabulated summary of what was observed in interviews one, two and three.

Int. 1	Male examiner. Student veiled, eyes showing (0 difference)					
	Paralinguistic/ non-linguistic events		Linguistic events		Accommodation observed <b>A</b> Non-accommodation observed <b>NA</b>	
	Examiner: neutral expression, occasional smile	Candidate: no facial expression visible	Examiner:	Candidate: no facial expression visible	Examiner:	Candidate:
Part 1 31:60 secs	Removes hand Direct gaze, head to one side	Hand on the table Lots of hand mvmt Gentle nodding of head Sitting upright			NA A	NA A
Part 2	Rh on hip Palms up	Hands tightly clasped Animated hand gestures Shrugs shoulders			A	A
Part 3	Both hands on table, rotating; sense of unease	Hands tightly clasped when not gesticulating			A –examiner trying to relax the candidate – trying to get the student to unclasp hands	

Int. 2	Female examiner. Student nonveiled. (0 difference)					
	Paralinguistic/ non-linguistic events		Linguistic events		Accommodation observed <b>A</b> Non-accommodation observed <b>NA</b>	
	<b>Examiner:</b> neutral expression, occasional smile	<b>Candidate:</b>	<b>Examiner:</b>	<b>Candidate:</b>	<b>Examiner:</b>	<b>Candidate:</b>
Part 1 31:60 secs	Direct gaze Smiley Leans in towards C	Direct gaze Smiley  Lots of hand mvmnt Sitting upright			A	A
Part 2	smiles	Fixing shayla Shrugs shoulders			A	A
Part 3	Hand mvmnts Slight swing on the chair	Hands mvmnts Crosses arms across body, clasping upper arms.	<i>What are the advantages and disadvantages of internet shopping?</i>	C clearly stressed by this Q	A – examiner mirroring C hand mvmnt & chair swing Relaxed stance, attentive & direct eye contact, regular smiles.	A – occasional nods Swing in chair

Int. 3	Female examiner. Student nonveiled. ( 0 difference)					
	Paralinguistic/ non-linguistic events		Linguistic events		Accommodation observed <b>A</b> Non-accommodation observed <b>NA</b>	
	<b>Examiner:</b> Open, warm and friendly expression	<b>Candidate:</b>	<b>Examiner:</b>	<b>Candidate:</b> no facial expression visible	<b>Examiner:</b>	<b>Candidate:</b>
Part 1 31:60 secs	Direct gaze, head to one side	Seated close in to the table Continuous gaze Frequent hand mvmnt			A	A
Part 2	Direct gaze Smiles nods	Direct gaze Smiles			A	A
Part 3	Nods Chair close in to table	Direct gaze Slightly swinging on chair			A	A

As can be seen in the tabulated summaries, the interviews were fairly uneventful in terms of observable differences in behavior other than those which could be explained by exam stress; nervous clasping of hands, sitting bolt upright for example. There are several exchanges which could be interpreted as incidences of accommodation; in particular the mirroring of hand gestures, swinging in the chair, direct gaze. In the first interview, there is an ‘event’ with the hand position of both: the examiner removes his as the candidates places hers on the table – she may have been trying to accommodate, he may have felt uncomfortable, culturally accommodating her as he would assume that unnecessarily close physical proximity

between the female candidate and the examiner would make the candidate more nervous or stressed. In the two interviews involving a female examiner, there are no perceived events of non-accommodation, or divergent behavior. There is clear mirroring of non-verbal communication in the nodding, smiling and direct gaze that both parties engage in. In these three interviews, there was no grade difference between the audio and face to face grades and that is significant when the application of CAT also reveals very little in terms of divergent or non-accommodative behaviour.

In interview 4, the candidate is fully veiled and with a male examiner.

Int. 4	Male examiner. Student completely veiled – no eye contact or facial expression visible (0.5 lower in audio)					
	Paralinguistic/ non-linguistic events		Linguistic events		Accommodation observed <span style="color: blue;">A</span> Non-accommodation observed <span style="color: red;">NA</span>	
	Examiner:	Candidate:	Examiner:	Candidate: no facial expression visible	Examiner:	Candidate:
Part 1 31:60 secs	Looks down Chair pushed away from desk <span style="color: red;">Restless</span>	Sitting at arm's length from table <span style="color: red;">Fiddles with veil</span>			<span style="color: red;">NA</span>	<span style="color: red;">NA</span>
Part 2	As C lifts veil to read card, A sits back sharply	Swinging in chair Rapid RH mvmnt			<span style="color: red;">NA</span>	<span style="color: red;">NA</span>
Part 3	Scratches leg <span style="color: red;">Fiddles with tie</span>  <span style="color: blue;">Gestures &amp; eye contact</span>	Hand on table <span style="color: blue;">Leans forward slightly</span>	Tell me about Arabic forms of Art		<span style="color: red;">NA</span> <span style="color: blue;">A</span>	<span style="color: blue;">A: tries to broach the gap in communication, physically leaning forward, hand on table</span>

As can be seen, initially the interaction seems to be marked by unease and discomfort; both parties neither mirror nor accommodate each other. The examiner does not look up at the candidate at all until the third part. The strategy is clearly divergent. At this stage, the candidate is struggling with what to say about Art (identified as a difficult topic by the students in the questionnaire) and puts her hand on the table and the examiner responds by looking up at her and gesturing the motion of writing, in this case, calligraphy. Each is accommodating and converging with the other. I would argue that it is this level of accommodation that results in the higher grade being awarded for the face to face interview – they are communicating successfully, without words and eye contact but with a sense of the other's unease trying to reach an understanding. This does not, however, enhance



the audio grade which remains a half band lower. It is possible that the face to face examiner is over-compensating, however, because the student is veiled and this has led to the higher face to face grade.

In interview 5, male examiner and female non-veiled candidate, the communication is uneventful in terms of paralinguistic expression. Eye contact is maintained throughout and both assessor and candidate appear to be relaxed. There is a minimal grade difference between the audio & face to face interview of 0.25.

Similar observations can be made about interviews seven and eleven which were between female examiners and non-veiled candidates. In these three interviews, it could be said that gender is not an issue affecting convergence or divergence strategies.

Int. 6	Male examiner. Student veiled –eyes visible ( 0.5 lower in audio)					
	Paralinguistic/ non-linguistic events		Linguistic events		Accommodation observed <b>A</b> Non-accommodation observed <b>NA</b>	
	Examiner:	Candidate:	Examiner:	Candidate: no facial expression visible	Examiner:	Candidate:
Part 1	Direct gaze	Sitting at arm's length from table Fiddles with veil	Do you prefer to get money or a gift? Do you listen to music often? Why don't you listen to music very often? Is it important in your culture?		A Hands waving side to side - pros & cons of music? Hand waves more when touching on a sensitive subject?	NA Initially distant, divergent, focusing on veil..for protection? Uncomfortable with topic of music?
Part 2	Sitting back in chair	Looks down when searching for vocab/ideas.			NA	NA averting gaze, diverting from the assessor
Part 3	Seated back, shoulders relaxed	Rh down towards floor - indicating sand on ground. Circular Mvmt of arm to show tyre/wheel Leans forward slightly	Can you describe to me the types of activities that children used to enjoy in the past.		NA A Nods & smiles but can't see if the smile is reciprocated.	A: tries to broach the gap in communication, physically acting out the game with a tyre played as a child

In interview six, the examiner tries to accommodate and converge by smiling and nodding to the candidate. His direct gaze is averted by the candidate when she is searching for vocabulary. He cannot see whether his smile is reciprocated but continues to smile and use direct eye contact throughout. He senses that 'music' is a sensitive topic, either from his own

knowledge of Islamic culture or from the student's clear unease (she begins to fiddle with her veil). She tries to converge, particularly in part three when she is trying to explain a childhood game that involved a car tyre; one that she feels needs further explaining via hand gestures and bodily movement. The candidate possibly feels the need to do this to accommodate the examiner into her cultural background. This was a common childhood game that the candidate may have thought would not be part of the examiner's schema, and therefore required further means to communicate this than she could achieve linguistically. Consequently, while the examiner may have understood the communication, the audio examiner would not have seen this dramatic 'extra' and this may have contributed to the lower (0.5) grade from him.

Interview eight was characterized by copious hand movements on the part of both assessor (F) and candidate (non-veiled).

Int. 8	Female examiner. Student non veiled ( 0.5 higher in audio)					
	Paralinguistic/ non-linguistic events		Linguistic events		Accommodation observed <b>A</b> Non-accommodation observed <b>NA</b>	
	Examiner:	Candidate:	Examiner:	Candidate:	Examiner:	Candidate:
Part 1	Direct gaze	Very noticeable hand mvmnts fixing shayla nods frequently looks directly at a but also shifts gaze frequently			A Eye contact maintained	NA shifting gaze, diversion strategy
Part 2	Smiling, leaning forward	direct gaze	Now I'm going to give you a topic: I'd like you to describe your favourite shop.		NA	NA
Part 3	V.Expressive hand mvmnts	Very noticeable hand mvmnts again Sits upright	Are there any other differences between now and how your grandparents shopped, for example?		NA A mirrors the lively hand movements	A: tries to broach the gap in communication, physically sitting up and gesturing with her hands

The examiner here works hard to converge with and accommodate the candidate with frequent smiles, direct eye contact and in part three mirroring the candidates expressive hand gestures. Significantly, the face to face grade is half a band lower perhaps indicating that the examiner was



aware that this level of paralinguistic communication was detracting from the linguistic communication and perhaps that without the gestures, the comprehensible communication was weak.

In interviews nine and ten, both assessors are male, with non-veiled female candidates and choose the same topics in part two/three; Hobbies.

Int. 9 & 10	Male examiner. Student non veiled ( 0.25 lower in audio)					
	Paralinguistic/ non-linguistic events		Linguistic events		Accommodation observed <b>A</b> Non-accommodation observed <b>NA</b>	
	Examiner:	Candidate:	Examiner:	Candidate:	Examiner:	Candidate:
Part 3 Int 9	Direct gaze Frequent head nods	Blinking a little and averting gaze Both hands moving Frequent hand movements	Do you think that the hobbies people do have changed?		A Eye contact maintained Nodding head as sign of approval/understanding	NA shifting gaze, diversion strategy
Part 3 Int 10	Playing with pencil, rotating & flicking Frequent nodding Smile	Fiddling with pencil under the desk Direct gaze Draws the hopscotch plan on paper	Can you describe to me the different types of hobbies people enjoy in your culture.		A Smiling Nodding head as sign of approval/understanding	A Mirroring fidgeting with pencil Direct gaze Drawing to facilitate understanding

In both cases the examiners try to converge with the candidates with warm friendly facial expression, nodding and direct eye contact. The candidate in interview nine appears to be uncomfortable with direct eye contact and tries to avert her gaze, whereas in interview ten, the candidate reciprocates. The topic of hobbies appears to cause both candidates to use paralinguistic communication to proceed: candidate nine uses hand gestures and candidate ten actually draws what she means on paper. The audio examiner, not able to witness this, scores both candidates a marginally lower grade. The concept of “hobbies” is in fact very culture-specific. In developing countries, it could be argued that there is no place for hobbies when most of the adult population is concerned with meeting their daily needs of food and shelter. Who then has time for hobbies? The UAE’s history, pre-oil, was one marked by a harsh existence for its inhabitants, who were largely poor and living off the fishing or pearling industry. The culture of hobbies was

therefore not part of their heritage and this could explain why both candidates had difficulty in answering the question.

#### Overview of face to face interview data

Applying a theoretical tool, such as CAT, is problematic in that the interpretation of the expressions of convergence and divergence are in themselves culture-bound. The observer is applying their own interpretation of what constitutes these features and in this case the features targeted were: eye movement, head movement, facial expression, posture, proxemics, hand movements, fidgeting and 'other'. The most noticeable expressions were via hand and head movements. In general, the proxemics were dictated by the position of the desk and chairs and so there was little variation. Facial gestures tended to be limited to smiles as the examiner's role is to remain neutral as much as possible and this was mirrored by the candidates largely. In contrast to what might have been expected, eye contact in most cases was direct and continuous. This goes against what is generally felt, that is that Muslim women do not feel comfortable with direct eye contact, especially with males (Parker, 1998). This was clearly not so from the evidence of this case study. If the grades alone are considered, the evidence suggests that paralinguistic communication indeed has an impact on grade outcome. For ten of the 11 interviews, the grades were the same (three) or lower (seven) for the audio recordings.

This has implications for both the examination process and teaching and learning strategies. If there were two assessors in the face to face exam, one could focus on the interview rubrics, the other in observing and grading the whole interaction, not simply the linguistic exchange. Secondly, teachers need to be made aware of the significant impact that paralinguistic communication has on the grade outcome and perhaps adapt their teaching of speaking skills accordingly.

### 3.3 Questionnaire Data

Two sets of questionnaires were administered, one to faculty and one to students. The main aim was to provide triangulation of the data and to find out a little more about the socio-cultural constructs framing the context and the interaction for the interviews.

#### Student questionnaires (see Appendix 6)

##### Section A: background information and language learning context

72 year two students completed questionnaires. All were Emirati female students, aged between 17 and 22, studying English and working towards their IELTS exam. The students came from a variety of program areas including Engineering, Education, Health, Business and I.T and most had had female, Arabic-speaking teachers in High School before coming to the college where they are taught by both males and females who are non-Arabs. 91.6% of the students' mother tongue is Arabic, the remaining 8.4% having Indian or Iranian mothers. The students speak other languages to a basic, good or advanced level but none considered themselves to be bilingual. The range of countries some have visited is broad, the most visited being Saudi Arabia (14), India (8), Thailand (8) and Iran (6). Many of the other countries listed had only been visited by one or two students. It is fair to say that the number of visits represents only three to five percent of the student sample. 69.4% of the students in the sample had teachers from the U.K or Ireland, with 18% from Australia or New Zealand and only 4% from America or Canada. 8.3% did not know where the teacher came from. So, the students in this study were exposed largely to a British/Irish and Australian/New Zealand delivery of English. The majority of the students, 77.7%, thought that it was very important to have a native speaker as their English teacher, unlike their High School experience. Many of the comments mention that having a native speaker will lead to them having better pronunciation and that they are forced to speak in English rather than Arabic. With regards their teachers, 73.6% felt that they had a basic understanding of Arabic and Arabic culture, although 18% did not and 8.3% were unsure. The questionnaire did not ask students to stipulate the names



of the teachers so it is possible that the 18% refers to the same teacher or it could be that perceptions of the teachers are not shared by the students.

When asked whether they thought it was important for teachers to be aware of Arabic culture and know a little of the language, 69.4% thought it was very important or important and the remainder thought it unimportant or were unsure. Comments in favour included: 'she can respect our values'; 'help us to understand each other'; 'she will know our culture and she will teach us equally with it'. Clearly for these students it is important for them that the English teacher shows empathy and understanding towards their language and culture: they are keen that the understanding be mutual. For those students who did not think it was important, their comments ranged from: 'it will not help him at all' to 'he does not need to know. He teach us English not Arabic'. So, on balance, more students wanted English teachers to be aware of their language and culture (by implication the religious practices of Islam too); they felt that mutual cultural understanding would facilitate their own language learning.

#### Section B: speaking assessment

Section B is specifically concerned with the questions asked during the assessment. There are certain topics which are considered generic and equally accessible and in part one, these topics are thought to be ones that the students can talk about freely, with ownership. The majority of the students (between 80 – 91.6%) were happy to talk about the topics of family, friends, themselves, home life, college life and free time activities. However, a significant number (between 8% and 15.2%) were not very willing and not at all willing to talk about free time activities. One student commented: 'I don't practice so much hobbies so I don't know what to say' and another wrote; 'I don't feel comfortable when I talk about my family'. An assumption has been made about what students may feel able to talk about freely and for up to 15%, this assumption is wrong. Question 15 is even more revealing. It asks students to state which topics they would feel

able to talk about in part 2 for up to 2 minutes. Results were much more scattered across the board.

Topics students were willing to talk about	Topics students were not willing to talk about
Weddings (87.5%)	International news and events (61.1%)
T.V programs (83.3%)	Flowers and plants (59.7%)
Travel abroad (79.1%)	Art (55.5%)
A memorable day (69.4%)	Famous buildings (52.7%)
Transport (68%)	Concerts and plays (51.3%)
Fashion (65.2%)	
Gift giving and receiving (63.8%)	
Celebrities (61.1%)	

For all of the topics there were students who did not know whether they could talk about the topic or not, the highest % being for flowers and plants (16.6%) and concerts and plays (15.2%). Comments from the students include: ‘because we don’t have knowledge on this’ (music); ‘I don’t have interest in these topics’ (books, flowers, news); ‘some of the students don’t listen to music because of our religion’ (music). It is possible that those topics that students are less willing to talk about are those that are not within their day to day experience. For example, going to concerts or plays is not something that Emirati students, especially females might do and interest in current affairs is generally low. Music is a hot potato as it is considered ‘haram’ (forbidden) in Islam to listen to music, although many of the students do listen to music in college.

With regards the assessment venue, the majority of the students felt that it was appropriate (80%), although comments were that the room was ‘tiny...like in a police station’ and one student said ‘I don’t like being alone with a strange man’. The overwhelming majority expressed no particular preference for the gender of the examiner (65%), but 27.7% would prefer a male examiner and only 6.9% a female examiner. Student comments in favour of a male examiner included: ‘The women are more strict and that’s

bad'; 'because female is very strong'; '...I think that some females are biased when they grade us'. 83% of students were happy with the 1:1 interview set up, 5.5% had no particular preference and 11% would prefer either a group interview or a 2:1 ratio so that 'if I stop, she can complete and help me to understand the question'.

With regards to wearing a veil, only 22% of the respondents said they preferred to wear a veil during the interview and 36% said that this decision was based on the gender of the interviewer. Two thirds of the students therefore were not influenced by the gender of the examiner in their decision to veil or not. Over half of the students, 68%, felt that wearing a veil would not affect the quality of communication during the assessment and, in fact, 8% thought that communication might be improved. On the other hand, only 11% felt that examiners would regard the veil positively and 65% felt that this regard would be either indifferent (34%) or negative (31%). When asked about the possibility of the examiner being veiled, the reaction was more positive (25%), with 85% unsure and the rest either indifferent or negative. When asked if they would prefer a non face to face interview (online/telephone) the response was in favour of face to face (57%), although almost 28% were in favour. Comments included: 'I feel more confident when I speak face to face'; 'to see the expression'; 'face to face is more personal'.

### Section C: varieties of English

Students were asked if they could distinguish between British, American & Australian varieties of English; over half (54%) said they could, compared with 25% who said they could not. Following on from that, students were asked if they thought they should speak in a particular way and if that would affect the grade outcome. 38% responded that it would make no difference, whereas 22% felt the British accent would be the best to use, compared with 14% favouring the American accent, the Australian accent trailed at 3%. 46% responded that examiners would not differentiate in terms of grade outcome depending on the accent spoken, but 35% were

unsure and 19% felt that examiners were looking for a particular accent; British (12%), American (5%), Australian (1%). Comments included: 'all English are the same', 'it depends for the examiner person and his thinking', 'yes British...because they are friendlier'.

### Faculty questionnaires (see Appendix 7)

#### Section A: background information and language learning context

Of the ten faculty/examiner respondents, four were female, six male and all aged 31 or above. All were 'mother tongue' English speakers from Britain (6), Ireland (2), Australia (2). The respondents spoke a variety of other languages to advanced level (none fluently) and between them had worked in 24 other countries. 60% felt that it was important, (40% very important) to have a knowledge of Arabic language and culture in the context of the classroom: 'I think it is very important to know the language to understand mistakes/errors students make, and the culture to understand which topics are better to use/not appropriate.'; 'not so much the language, but understanding of the culture/tradition is very important for rapport'. 70% felt that the students appreciate their teacher having knowledge of Arabic language and culture: 'I think it reassures them'; 'again for rapport'; they like the 'respect' aspect of knowing about them/their culture/language. 70% of the teachers felt that their students value a native speaker. One of the teachers commented that there is 'still a notion of correct English-Western native speaker'; in many countries, students expect a white native English speaker. Otherwise, they feel they are not getting the 'real deal'.

#### Section B: speaking assessment

On the whole, teachers felt that all of the topics in part 1 were met with favourably by students. Four teachers expressed slight concern over the topics of free time activities (2), family (1) and friends (1). With regards the part 2 topics, the topics that, according to the examiners, were the most and least popular are shown in the table below.

Topics examiners felt students were willing to talk about	Topics examiners felt students were not as willing to talk about
Weddings (90%)	Art (70%)
T.V programs (80%)	Concerts and plays (50%)
Travel abroad (79.1%)	International news (40%)
Fashion (70%)	Music (40%)
A memorable day (70%)	
Religious festivals (70%)	

Most (70%) felt that the interview room was an appropriate context but opinion was divided with regards to the gender of the examiner. 50% felt that students would prefer a female examiner compared with male (10%) and 40% being unsure or recording 'no preference'. With regards to students wearing the veil, 50% felt that this was not appropriate either partially or fully, for the interview. None felt that it was appropriate for an examiner to be veiled, 70% definitely, 30% unsure. Only 30% felt that if the student wore a veil, her quality of communication would not be affected, compared with 50% who felt communication would be impaired. In fact, 40% felt that the grade outcome would be worse where the student was veiled, 30% were unsure and 30% felt that there would be no impact on the grades. When asked about the examiner being veiled, the examiners were clearly unsure of the impact (70%) but some felt (20%) that the students would react negatively. The response was almost unanimous in regard to non face to face interviews: 90% felt that online interviews would not work, 10% were unsure.

### Section C: varieties of English

All of the examiner respondents felt able to distinguish between varieties of English and 60% felt that it is not important how they speak during the exam, whilst 20% said British English was what should be spoken. With regard to the students, all felt that it did not matter which variety of English the candidate spoke.

### Overview of questionnaire data

When we cross-reference the data from the questionnaires, the results reveal clear areas where opinions and attitudes from students (candidates) and faculty (examiners) converge and diverge. These attitudes and opinions inform and can shape the interaction that takes place during the interview in the form of accommodation or non-accommodation.

The main areas of convergence are in the areas of native English speaking examiners and attitudes to Arabic language and Arabic culture. Both parties agree that students see native speakers as very important to them and that these teacher/examiners should have a basic knowledge of Arabic language and culture. The students identified the same topics as being accessible in part 1 as the examiners; family, friends and home life. They also agreed that in part two, the topics of weddings, T.V , a memorable day and fashion would be good topics and that they would be less willing to talk about international news, concerts and art. The venue seemed acceptable by all for the interview.

Areas of divergence included topics in part two, questions regarding the gender of the examiner, the wearing of the veil and face to face exams. In part two, students identified the topics of travel, transport, gifts and celebrities as being topics they would be willing to talk about and did not identify the topic of religious festivals, contrary to feedback from the examiners. Students highlighted the topics of flowers and famous buildings as ones they would not be willing to talk about, the examiners did not seem aware of this according to their responses. With regard to the gender of the examiner, contrary to the examiner's suppositions (50%), only 6% of the students responded positively in favour of a female examiner and 67% of students, compared with 30% of examiners, felt that wearing a veil during the exam would have no impact on the final grade. The other area of significant divergence is the question of face to face interviews compared with online interviews. Clearly this is a hypothetical question (assessed interviews have never been conducted this way before in the college) and

therefore the responses are only valid in respect of that context, however, 90% of the examiners felt that face to face interviews were preferable to online interviews whereas only 57% of the students felt this way. The key areas of divergence as evident from the questionnaire data are therefore topic choices, the impact of the veil on the grade outcome and the gender of the examiner.

3.4 Focus Group Data (see Appendices 8 and 9)

Two sets of focus groups took place: one for examiners and two of students. The students were asked about their experiences having just completed mock IELTS-style interviews. One student had difficulty speaking about sports in part one:

I don't practice sports...
And I don't know anything about it.

Another student had difficulty in part two:

When I said about er my favorite transportation, he kept asking me about the transportation itself, then the people, what are the problems, what are the benefits.
And what are you expect in the future.
All of these, some of them I know how, how to answer.
But not all of them.
I was like sometimes stuck/stop.
I don't know what to say.
Because of I do not have experience and er maybe I didn't have much knowledge about this thing.

Clearly the topics are of major concern to the students: they feel (see above) that they are hampered not by ability to speak but by “experience” of the topic area. They do not seem to be worried by the gender of the examiner, however. When one student was asked whether the interviewer (male) allowed her to show off her speaking skills she replied:

Yeah. Because when they looked at er us and by eyes,eyes contact,
and they smiled.
They make us feel comfortable when we talked.

The assumption is often made that female candidates are uncomfortable with direct eye contact from the examiners, but this was not the case here. When asked what changes they would like to make to any part of the assessment, students' responses were squared firmly on topic choices. They suggested that they should be able to choose their topics and that the fact that part two ran into part three (extension of the topic in part two) meant that if they struggled in part two with the topic, then they were set up to fail in part three:

Yeah. Because the part, er three is related to part two.
If we don't know to speak in part two so we will er kinda.
(confus in part three)

They would prefer to see topics related to their own society:

Er...
Maybe something about our society's issue.
Ehm...
The ,the enormous increase in divorce rate.
The, the unemployment.

The students recognize the need for proficiency in English in relation to many aspects of their lives such as in the wider community (shops/hospitals etc), when they travel abroad, to communicate with their housemaids and drivers, to communicate at college and most importantly for the world of work:

Because in our, in our job it will help us to communicate with other peoples.
And because er er speaking is er er a main part, I think so.
Bacause er in gen, in a business or er and it's now English becomes
er er our language.
So we have to, speak
Yes.
Travelling around the world.
Yeah. Because sometimes just as you have to speak in English.
They don't take Arabic language.
So it's easyto communicate in English.



In our job also they, we must have ENGLISH to access this job nowadays.
-------------------------------------------------------------------------

Faculty focus groups

The faculty also highlighted concerns about topic areas. One examiner mentioned that one of the questions they had to ask was “did you learn to drive whilst at school?” and another, ‘would you like to be a train driver?’. The first question is not sensible – how many schools teach their pupils to drive and the second is inappropriate since there are no trains in the Emirates; only recently was a Metro service opened in Dubai. Another topic queried is that of attending public events:

..in parts two and three. Because there's topic which you know they
might not be able to talk about like this attending public events.
I think they were looking for concerts or a carnival or possibly
football matches, something.
But especially the girls they don't, they don't seem to know what
To talk about.
They don't go to public events for them it's a big crime.
And that 's in the part two, and the part three

This is something that concerned the students too (see above p.99); if they are struggling with the topic in part two, they are also going to have difficulties in part 3 as they are interlinked.

In relation to part one, examiners ask whether it is wise to switch from topic to topic when the purpose is to relax the students and give them familiar topics to talk about. Just as there are queries about the links between parts two and three causing potential difficulty for the students, part one where examiners must ask about at least two other topics is seen as problematic.

Another thing is in first part which Is supposed to help them relax
Not having complete shift in the topics.
..from.. Describe your neighborhood.
Let's talk about transport....

One of the examiners draws attention to the difficulty in assessing pronunciation in the interview;

differently would I, evaluate somebody because they struggled a
little bit just because I'm unfamiliar with their...
...with their accent.
With their, their language actually.
..people from the Far East they tested them.

But they are much more difficult .
Their accent is much more difficult.
It could be impenetrable.
Their Arabic accents..... in English is not that bad

It seems that if you have been acclimatised to a particular accent because of where you live and work, you may feel less strain in understanding the speaker than if you had never come across that accent before. This would have an impact on the grade you would give for pronunciation. Similarly, one examiner mentions vocabulary items that will be familiar to some and not others depending upon their experience as teachers and examiners around the world:

...but I have a thing with your accomodation like...in
speaking they always doing, talking about food most especially
actually talking about 'harees'.

Teacher/examiners in the Gulf would know that 'harees' is a Gulf dish that is served during Ramadan, for which there is no translation; it is simply a name. The examiner is suggesting here that knowledge of local culture and society does have an impact on the way examiners may grade.

Another issue is that of making eye contact with students. Examiners (especially male with female candidates) are unsure whether or not to gaze directly at female candidates or, as pointed out below whether it will be seen as rude if they avert their gaze.

Yeah. Students complaining that you know that, the examiner were
asked the questions and then some cases they look out the window.

Sensitivity to the candidates' nationality is also an issue highlighted by one of the examiners. One of the questions in part one is often about the home area:

What about those people from Baghdad and the one...tell me about
where you grew up.
Oh.Yeah...
It's not a place to grow up.....
And what would you tell visitors to do there?
What would you recommend....

Apart from the choice of topics, the examiners are also concerned about procedural issues and one recommends that perhaps there should be two examiners as it is very difficult to ensure you are following the rubrics and listening to the candidate at the same time:

I think like speaking exams having two people
a talker and a grader.
I'm sometimes just so worried about procedures and timing.
You do lose track of what they say ....

### Overview of focus group data

The focus group data reveals that the students and examiners have similar areas of concern about the range and suitability of the topics available in the assessment. They also share a concern about procedural issues; the students would like to be able to choose their topics in part two and the examiners would rather there be two examiners, an interlocutor and an assessor for the exam. Examiners are also concerned about cultural appropriacy in terms of questions (ref. Baghdad) and acculturation towards language items (ref. harees).

## **CHAPTER 4: DISCUSSION and CONCLUSIONS**

This case study is a snap shot, a vignette, of a small group of teachers and students in a very specific context. However, the results from the research uncover some issues that are generic and of interest to a much broader audience, locally, regionally and internationally. The evidence suggests that there is a clear difference between what is 'experienced' (face to face) in an interview and what is 'heard' (audio) and that these differences can have a direct impact on a student's grade outcome. This is of universal importance to everyone involved in the EFL assessment process, all stakeholders from the test designers to the administrators to the examiners, teachers and candidates.

Additionally, the nature of the communication is affected by the sociohistorical background of the interactants. Native speakers, for example, tend to develop topics as an information exchange, based on shared meaning and understanding, whereas non-native speakers are less likely to share as much meaning and understanding. This is borne out by the case study: speech is an act of co-constructing shared meaning based on shared understanding. Interaction is more effective when both parties are from the same sociocultural and linguistic background. In the case of international English assessment interviews this balance is compromised because of the difference in schemata that the two participants bring with them to the interview. It is suggested that the extent to which convergence strategies are used can have an impact on grade outcome, and CAT is the tool that is employed to ascertain the degree and impact of accommodation. The differences in schemas and the degree of accommodation that takes place can both affect the reliability of the assessment tool.

Research by Berwick and Ross (1996) suggests that where reliability may differ is if examiners accommodate (or use divergent strategies) to the candidate based on their own cultural background. Brown and Hill (2007a) studied interviewer style and its affect on grade outcome in IELTS interview

assessments and the results showed differences in grades for fluency and pronunciation (this is replicated in this case study) and clearly indicated that the extent to which an assessor develops topics, reacts to the candidate's responses, in short, all examiner behaviour has an impact on the language produced by the candidate. Interestingly, one of the examiners was an IELTS trained assessor and the candidate scored a lower grade in this case. Nakatsuhara (2008) suggests that this could be because of the standardization training that IELTS examiners undergo. That implication is that assessor training can reduce the difference in assessor behaviour, but to what extent does this then invalidate or 'corrupt' the communication that takes place? If the input from the assessor is scripted and, 'rehearsed', to what extent can the candidate's input, as co-constructor of the communication, be assessable? Their response is effectively to a scripted speech rather than a 'live' input with all its spontaneity and unpredictability. This raises the question of whether it is valid to assess a candidate's ability solely on their performance when that performance is framed by so many other variables: the test design; the examiner; the rater (if not the same as the examiner); the rating framework and the physical context of the test (McNamara 2000 p.21) .

This final chapter will consider the following:

- 1. Research questions and discussion of results
- 2. Research methodology reviewed
- 3. Research findings
- 4. Conclusions
- 5. Opportunities for further research

#### 4.1. Research questions and discussion of results

The over-arching question in this research concerned the extent to which international EFL oral assessments could be universal and still remain valid and accurate measuring tools in any context. Within that question, several other sub-questions emerged, including:

- 1. What role does ‘accommodation’ play in the negotiation of meaning in oral assessment?
- 2. To what extent can the IELTS publically-available marking criteria really reflect true communicative ability?
- 3. To what extent do the examiner’s choice of topic and the student’s ‘world knowledge’ help or hinder their performance in the test?
- 4. To what extent does the wearing of a veil impact the perceived quality of communication and the grade outcome?
- 5. Is examiner gender, in this context, a factor affecting candidate grade outcome?

In this particular context, that is an all-female student campus with a restricted intake (Emirati females only, no other nationalities) the research questions apply, very specifically, to a controlled situation. Is it important that they are seen in this light because quite different results may have been found even with an all-female campus but with not all the students being of the same nationality or from the same cultural background. It is that very background, however, that highlights some of the issues listed above and is instrumental in shedding light on the nature of intercultural communication between a homogenously ethnic group (the students) and a more heterogeneous ethnic group (the examiners).

Three of the sub-questions relate specifically to the cultural background of the students; cultural background, topic choice, wearing the veil and world knowledge. The other two questions have more of a generic application, although cultural background as a factor can be applied to other cultures, not just the one highlighted in the case study. The effect of gender on candidate outcome (both examiner and candidate) is a factor that could be investigated in any context as is the applicability of the IELTS grading system and the extent to which different varieties of English are considered acceptable in the assessment of spoken English.

All of the sub-questions could have been addressed as a whole research project in their own right, but within the scope of this case study, that would

not have been possible. The aim of the case study was therefore to look at the assessment process as a whole, from within this very specific cultural context, and to try to relate it to some of the broader issues and concerns related to EFL speaking assessments.

#### 4.2. Research methodology

The research methodology is deliberately interpretivist in nature. As mentioned above, it is a case study socially constructed by the researcher in that the participants were gathered by the researcher within a very specific context. The interpretation of the construction of the language and meaning of that context can be viewed differently according to the agent(s) of that interpretation.

The methodology chosen, of filmed interviews, focus groups and questionnaires generated a lot of data in relation to the research question and sub-questions. The data presented some very interesting findings which challenge assumptions, particularly about the sociocultural context of the research.

The filming and audio recording of the interviews was semi-constructed in that the participants came from the same college campus but those participants were not chosen for any specific qualities other than they were year 2 students and English faculty with IELTS examiner status. They were all volunteers and no volunteers were rejected; there was no pre-selection process. The matching of candidate to examiner was purely on the grounds of availability according to timetable constraints. This part of the methodology was central to the research because its intention was to highlight i) the complex nature of assessing spoken language and further, within the restrictions of the given context and ii) to reveal the difference in nature of a face to face interview from an assessor's point of view, to that of an audio recording.

The chosen tool for analysis of the data was CAT because it allowed for the interpretive nature of the case study whilst generating solid, grounded data.

The assumptions of CAT are particularly relevant to the research questions because CAT acknowledges the sociocultural context of individuals and the fact that communication is not simply about speech or linguistic behavior but that it is about paralinguistic, discursive, and non-linguistic moves between interactants. Moreover, CAT acknowledges that these non-linguistic features are equally important. This challenges the very core of EFL assessment, where there is no acknowledgement of this aspect of communication and its possible impact on a grade outcome in the grading criteria. The theory of CAT was perceived by the researcher, therefore, as a perfect tool, however the practical application of CAT was problematic. One of the issues, which was not dealt with in this study, was the extent to which any convergence or divergence that was perceived to be happening was a conscious move or not, as Thakerar et al (1982 p.247) commented 'the measured linguistic shifts of convergence and divergence need not necessarily reflect the direction in which the speakers themselves believe they are moving.

Furthermore, there is the assumption in CAT that accommodation is motivated by the speaker seeking approval: in the case of EFL interviews this is evident in that candidates try to align their speech to that of the perceived notion of what the examiner is expecting to hear, thereby seeking approval and hopefully an appropriate grade outcome. From the questionnaire data, it is evident that the 'type' or variety of English is not a major consideration for either the candidate or the examiner: standard American or British English is not consciously striven for by the students to achieve success. The overriding motivation for the candidate is to maximise grade outcome and variety of English is only one consideration. Other strategies used, both consciously and unconsciously, by examiner and candidate help in the negotiation and the co-construction of the communication.

There was little previous research data available on the use of CAT in this particular application, that is, in language assessment interviews. In fact,



CAT has been subject to criticism because of the fact that the original theorists behind it were not linguists themselves (Trudgill, 1994). Most of the research carried out has been in the fields of medicine, patient:doctor relationships; crime, police:criminal interrogation methods or social work. This research honed in on particular personality aspects of communication such as perceived trust or honesty. The present research in this case study was not specifically targeting one aspect of communication, but rather the whole communication process: paralinguistic and linguistic aspects. This presented some problems. With regard to the linguistic 'competencies', the band descriptors provided a measuring tool but for the paralinguistic competencies, CAT only really provided a theory without a practical application. Therefore it was necessary to develop a means of applying CAT to the data to be analysed.

To apply CAT to the research project, specific features of non-linguistic behavior were identified by the researcher and these were used as the yardsticks for measuring the extent to which convergence and divergence was occurring. The recording of these 'non-linguistic events' was carried out by the researcher and another independent examiner, and so must be regarded as potentially biased, or at least, coloured by, in the words of the CAT definition, the researcher's own 'personal and social identity'.

The observations and data recorded as grade outcomes for both face to face and recorded interviews provided rich data for analysis, but this could have been further enhanced by a third dimension. It would have been interesting to conduct a quick questionnaire or interview with the participants immediately after the grading process. In the case of the interview, it would have been useful to tap into 'the moment' immediately and ask both the examiner and the candidate for their impressions of the interview; did either of them feel uncomfortable with the room, the questions asked, the non-linguistic behaviour of the other, the rubric of the test and so on. This would have been interesting data because it would have captured the immediacy of the interaction without the 'regrouping of thoughts' that can happen when

one reflects on an experience after a prolonged time delay. This data would have had direct relevance to the interpretation of the perceived interaction. The use of CAT to analyse the interaction was that of an observer, second-hand, and it would have been extremely useful to have recorded this extra dimension from the participants themselves. Furthermore, this may have shed light as to how true to their usual behaviours was that of the participants in the interview or were they were acting so as to appear model candidates and examiners or behaving as they felt was required by the researcher.

The interview itself is a staged event and because it is so heavily scripted, it is not an authentic conversation: both parties are, in a sense, acting. It would have been interesting to find out whether the candidates and examiners themselves felt that they had exhibited different behavioural patterns or exaggerated or played down particular expressions during the interview. For instance, a student may have been told not to wave their hands around too much, the reason given that this is very distracting for the examiner. Hand waving for this student, however, may normally be an integral part of her communication strategy and therefore by asking her to restrain her hand movements, her speech delivery might be handicapped or reduced. At the very least, she may feel uncomfortable during the interview and very self-conscious about making any kind of hand movements. Add to this the extra dimension of a camera and this can accentuate self-consciousness further. Perhaps if cameras were the norm in interview contexts, then the data would be more reliable as the participants would not be so conscious of their presence.

The questionnaire data, for both examiners and student participants again was constructed by the researcher who directed the questioning process through a series of mostly multiple choice questions. The data must be seen as biased in the sense that the researcher led the train of thought but it is hoped that with the opportunity for respondents to add comments, this was counter-balanced to some extent. The data received from the questionnaires

was much more useful than that of the pilot project questionnaires and this was because more specific questions were asked, particularly in relation to topic choices and perceived difficulty of those topics.

The focus groups were an opportunity for the participants to hear about the early research findings and for them to offer their input and interpretations. The dialogue was only loosely guided, and whilst this was deliberate in order to allow the participants an open forum, it did mean that specific questions that the researcher had intended to pose, went unasked. The focus groups could have been followed up with the remaining questions presented in a group email, but it was acknowledged that the participants had already given very generously of their time and so it was decided not to follow this through. Time and logistical issues concerning timetables, meeting arrangements and the nature of the focus group feedback made it the most difficult part of the research triangulation to administer. Arranging for single interviews was relatively easy, but trying to co-ordinate all of the teachers and participants at their respective meeting points was problematic. Apart from logistical issues, there were other potential issues with the data in the sense that despite assuring the participants of the anonymity of the feedback, there is always the possibility that some comments are made because the respondent feels that is what the researcher wants to hear or is what they ought to be saying.

#### 4.3 Research findings

The research findings present a variety of significant observations relating to the assessment process itself and to the linguistic and non-linguistic interaction between both examiner and candidate. The key findings in relation to the research questions are with respect to:

- The measure of ‘accommodation’ in the negotiation of meaning in oral assessment.
- The extent to which the IELTS publically-available marking criteria can effectively assess communicative ability.

- The extent to which the student's 'world knowledge' and topic choice impacts 'performance' in the exam.
- The extent to which wearing a veil impacts the perceived quality of communication and the grade outcome?
- Examiner gender as a factor affecting candidate grade outcome.

In an IELTS test, as in the case study, the exam questions are scripted to avoid unpredictability and variance in question difficulty but this in turn destroys the very nature of communication in that it constrains the impromptu digressions that often occur in natural interactions. Even though the test itself is scripted for the examiner, clearly the delivery of the questions will vary from person to person. Brown, in a study of interviewer variation and the co-construction of speaking proficiency, concluded that the interviewer is 'intimately implicated in the construction of candidate proficiency' (Brown, A. 2007a). Various factors come in to play such as the perceived friendliness of the interviewer and level of rapport that is established; the choice of questions made by the examiner and their perceived easiness or difficulty; male/female dynamics and intercultural compatibility. All these factors were theorized upon via CAT.

#### 4.3.1. The measure of accommodation

CAT is a heuristic tool that in essence is quite simplistic. It offers the means to reflect on the mechanics of communication without being over-prescriptive. It is equally applicable to a variety of contexts and groups of interactants and thus suitable for looking at intercultural communication. Accommodation (Gudykunst, 2005 p. 137) is the 'process through which interactants regulate their communication'. CAT offers a means of observing that regulatory process which is manifested in convergence or divergence in both linguistic and behavioural output and is context dependent. The benefits of CAT are that its scope is wide-ranging, but as Gallois, Ogay and Giles point out (cited in Gudykunst, 2005 p138), it is expected that the researcher identify the markers that would define accommodation and non-accommodation. This was one of the

disadvantages, however. It meant that the adherence to a particular set of criteria was not required and the researcher could interpret the norms and set the boundaries according to the needs of the case study. On the one hand it was quite liberating, on the other, daunting. It was decided to focus on just a few observable non-linguistic markers such as frequency and duration of eye contact, head nodding and proxemics, for example. This freedom of choice, whilst welcome, did however open the possibility of researcher bias. The researcher, in setting the norms as measuring tools, had a critical perspective with its own biases. The norms identified were those perceived on the basis of the researcher's own sociocultural background, as shared by the participants, and must be considered in that light.

CAT was chosen but there are other theories such as Identity Management Theory (Imahori & Cupach), Identity Negotiation Theory (Ting-Toomey) and cultural Identity Theory (Collier & Thomas) which are all concerned with the mechanics of intercultural communication. CAT unlike the others mentioned (see the literature review for more details) seemed to be more all-embracing and offered the flexibility to set the norms independently.

CAT did inform and guide the observations and to that extent it was successful in illuminating some of the features of the co-construction of meaning in the oral assessments. In Gregory and Webster's (1996) study of the Larry King interviews, they found that perceived levels of status and power influenced the accommodation that took place. This was, in a sense replicated in the case study. In interview four, for example, the male examiner, who arguably holds more of the power balance in the context of the interview, does little to accommodate the veiled student's attempts to encourage eye contact. The motivations here though, were not for the examiner to present himself as having more power or influence, (in contrast perhaps to Larry King's guests) but the effect was the same: communication was made more difficult and this was evident in the audio grading. It was half a band lower than the face-to-face grade. The higher face-to-face grade suggest perhaps that the examiner was aware of the fact that they had not

converged non-verbally with the candidate, avoiding eye contact and then over-compensated through the grading. Where communication is facilitated and both parties appear to be converging, as in Gallois and Watson's study of health professionals (1998), the co-construction of meaning is arrived at more easily. In the case study this is borne out by the grades, as was seen in interviews where there were no differences in the face-to-face grades and the audio grades: interviews one, two and three. In all three interviews, the face-to-face interaction is marked by an effort on the part of the examiner to use facial gestures, hand movements and bodily stance to converge with the candidate. There were also several instances where students and examiners took advantage of employing non-verbal strategies, such as waving a hand over the shoulder to denote 'in the past'. This correlates with Jenkins and Parra's study (2003) and highlights the need for teachers to raise awareness about the role of non-verbal behavior such as mirroring and eye contact to their students as a tool to supplement their linguistic communication skills.

The theory needs to be further developed. From a researcher's standpoint, it was difficult to apply because it did not offer any methodological guidelines. Without a suggested structure and methodology of use, the theory risks being applied in many different ways and in a huge variety of contexts making generic deductions very difficult. Although to date used largely in the fields of health and crime, CAT provided the theory that could be successfully applied to the context of EFL oral assessments but with little information as to the means of applying it. As CAT is used more and the research documented, it is envisaged that a methodology will emerge that can be applied to different contexts. In this context, CAT revealed that convergent strategies did result in better communication in terms of successful co-construction of meaning.

### The measure of linguistic competency.

CAT highlighted and helped quantify the mechanics of co-construction of discourse and the communicative event in the face to face interviews. The data from the grading of the face to face interviews and the audio recordings clearly showed the importance and influence on the outcome of an exchange that real-time involvement has. In the audio recordings the examiner was a passive observer and was unable to witness the nonlinguistic interchange, relying solely on voice and speech. In contrast, the examiner in the interview was an active participant in the exchange. The resultant grade differences can be seen to prove the assumption that non-linguistic behaviour is an important factor in the communication process, however, the differences were small and could also be interpreted as normal margins of grade differences. Currently, the IELTS publically available Bands only assess linguistic ability and it has been clearly demonstrated that this is only one aspect of communication. In that sense the grading criteria fail both the examiners and the candidates as the data used for assessment purposes is only one-dimensional. Indeed it raises the question as to whether international speaking assessments should be videoed.

According to the data, grammar and pronunciation are marked consistently higher by the face to face examiners. This would seem to imply that those two linguistic features are definitely supported by non-linguistic behavior, since the absence of its observation results in a lower grade by the audio examiner. Again, this is a significant finding, and perhaps unexpected. Pronunciation and grammar are usually considered to be more 'scientific' and countable features of linguistic communication, especially grammar. It is as though when face to face, the grammar is only part of the message being delivered and received: the degree to which it is correct or not is not recognized in the same way as it is in the vacuum that is audio grading. This evidence seems to suggest that the whole experience of the interview, with the examiner and candidate face to face is much more than the sum of all the parts as witnessed by an audio examiner. We know that fluency and

lexis can be expressed with non-linguistic behaviour. In one of the interviews, a student cannot remember the phrase 'in the past' and so demonstrates the passing of time by waving her hand over her shoulder. In the same way, fluency can be perceived, in part, as the extent to which the candidate interacts with her examiner both verbally and non-verbally. In two interviews, interview four and six, the grades given for fluency are each a band higher for the face to face interview and, significantly the face to face interviewer is male with a veiled and partially veiled candidate. Even with a veiled candidate, fluency in a face to face interview is graded higher than via the audio grading. Furthermore, this also suggests that the two male examiners, in this case, were perhaps accommodating or maybe trying to compensate for the lack of visual communication signals (due to the veils) and possibly due to the female:male counterpoint, and that this is the part of the process that is absent in audio grading. This points to the 'whole experience' of the interview and indicates that audio grading, as in IELTS exams, does not capture the communication event in full.

In interviews four, six and seven, the grades for grammar were higher in face to face grading (6:5, 6:5, 6:5 respectively). It could therefore be postulated that the perception of grammatical accuracy is equally subject to the forces of accommodation and non-linguistic behaviour as fluency and lexis. Again, interviews four and six were male examiners, this is significant because it further points to the fact that accommodation can take place even with veiled candidates and across genders in a linguistic area that might normally be perceived as unaffected by non-linguistic behavior. In interview seven, the face to face interviewer was female and the student was non-veiled and the results are similar. This would suggest that, irrespective of the gender of the interviewer, the perception of grammatical accuracy is affected by the nature of a face to face interview rather than just simply the tangible, measurable grammatical output from the candidate that is measured in the audio recording.



Pronunciation was the category that was the most contested by the examiners in the case study. Nakatsuhara's study (2008) concludes that students can be unfairly disadvantaged in the pronunciation category if the topic choice is outside their field of vocabulary, so grade outcome in pronunciation is directly linked to lexis and topic choice.

The examiners in this case study were from different English-speaking European countries and had their own individual accents; if this was amplified internationally, we would see an army of examiners each from a different English-speaking background and potentially with different perceptions of what constitutes correct pronunciation. James Dean Brown (2004) goes further suggesting that it is not only the type of English itself that could be considered a source of bias but also lists administration procedures; test directions; test content; test knowledge selection; testing method; rating/scoring; score interpretation, and norm sample selection. He also comments on the fact that within any test situation, there are a variety of Englishes at play, for example the English of the test content, of the test proctors, of the raters/scores etc and so who decides as to what constitutes English proficiency? According to the data of the research, again the face to face scores were consistently higher than the audio recording which, as with the grammar, seems to suggest that not only lexis and fluency (which are perhaps more obvious) but also pronunciation is positively affected by non-linguistic behaviour. Or is it simply that the examiner is fully engaged in the conversation with the candidate such that his/her ability to isolate discrete areas such as grammar and pronunciation is compromised? Perhaps it is impossible to focus on one language feature when you are engaged in the whole experience as interlocutor and examiner at one time and perhaps the examiner focuses less on grammar and pronunciation in a face to face interview because they may have less impact on the construction of meaning than lexis and fluency.

The conclusion is demonstrable. The skill of speaking is only one part of the many skills involved in effective communication. Speaking involves one person, but communication dictates that that one person's message has

been conveyed and understood: it is this partnership with the recipient, who plays an active part with the speaker that co-constructs the meaning of the message. An audio examiner is only a passive observer of the spoken word and therefore can only offer an assessment of one aspect of the test. Spoken words, in isolation, are not communication. It can be argued that a writer communicates with the reader using speech in isolation, however it is the punctuation, the choice of vocabulary, the sentence length and context of the writing that gives shape and meaning to the communication and helps co-construct that meaning between the author and reader. In the same way, the construction of meaning from the spoken word is not simply about the words themselves.

According to the IELTS “Official practice materials” booklet (2007), ‘the speaking assessment assesses whether candidates can communicate effectively in English’. This is true of course of the actual test conducted by IELTS, but where there are queries over a grade given, by a candidate, or where the examiner is being monitored, this is done via audio grading alone. The evidence points to the fact that this is not a valid means of moderation. This can only be achieved by either having an assessor and an interlocutor in the exam room, one to conduct the interview and one silent observer grading the candidate, or by videoing all the interviews.

Another possible alternative to videoing, which brings with it further potential problems such as objections to being filmed (examiner and candidate), might be to conduct online interviews via a webcam or videoconferencing media.

Developments in speech technology have facilitated the use of instructional software and assessment tools and supporters of it suggest that the grade outcomes are more reliable because human subjectivity is removed from the grading process. Bejar, (2010 p.4), whilst encouraging the use of technology in automated oral testing, also acknowledges that technology should be seen as an enabler, rather than a solution. A universal problem for both online and face to face tests is what criteria to use. The first task in creating the software tool is to compile a speech corpus that will inform the

speech-recognition engine. The speech corpus is problematic in the same way that the banding criteria used by human examiners is. Decisions have to be made about definitions of proficiency in terms of fluency and pronunciation, for example. The speech recognition software also has to be programmed to analyze and accept specified speech features. The scope of the software would have to be very broad in order to allow for the widest range of possible accents and be able to respond appropriately. As Bejar (2010 p.5) points out:

...speech recognizes need to exhibit higher recognition accuracy, which is a significant challenge given the nature of the speech being produced by non-native speakers of potentially many different native languages and proficiency levels.

Added to the limitations of the software, there is also the problem of potential technical break-downs, which would require technical support to fix. Human examiners can adapt to changing circumstances but when a computer system is down, the assessment cannot continue. Furthermore, both candidate and examiner would need to feel comfortable with the medium and be assured of privacy protection. As already stated, the interview is a 'whole experience' and part of that experience is that it is a live, real time exchange of speech and non-linguistic behaviour.

However, other studies point to the fact that online teaching and testing cannot simply be viewed as substitutes for these face-to face settings because meanings are constructed using different modes and media (Hampel and Hauck, 2006). The new context of an online environment presents a challenge because users of the technology must become comfortable and conversant with the new mode of communication and their responses to the media. Negotiating shared meaning through a different media adds a new dimension to communication and so Virtual testing is an important area for further investigation.

### Topic choice with regard to cultural context

There is evidence of cultural bias, or rather cultural misunderstanding or even misplaced cultural sensitivity in EFL assessments, in this case oral assessments. Topic choice in the assessments is one area that would seem to support this view. There is a clear mis-match between what students perceive as accessible topics to talk about and those that the examiner and test creators perceive to be reasonably generic to all nationalities and cultures. Khan (2005) in the University of Dhaka, found that IELTS questions 'assumed background knowledge and vocabulary beyond their (the candidates') range and exposure'. For instance, abstract topics, such as 'humour, peace and relics' proved challenging to some candidates, not only in terms of generating ideas but also as Nakatsuhara discovered, in terms of pronunciation and lexis. Topic choice, therefore, has a significant impact on possible grade outcome. Candidates in Khan's study also had difficulty in expressing their own opinion as 'they were restricted by their view that you shouldn't disagree with a teacher (examiner). Khan identified topics and vocabulary items in IELTS that reflected, she felt, western concepts, schema and patterns of interactions that effectively disadvantaged her students (Khan: 2005). Her students had difficulty in speaking about holidays, for example, a seemingly innocuous and generic topic, but as it was not in their schemas, their field of experience, this negatively affected their ability to offer evidence of their linguistic competencies in the exam. In this case study, topic choice was the area of greatest concern for students. There is evidence of divergence between the candidates' and examiners' perceptions of what would be an accessible topic

<b>Topics STUDENTS felt comfortable talking about</b>	<b>Topics EXAMINERS felt students were comfortable talking about</b>
87.5% Weddings	90% Weddings
83.3% T.V programs	80% T.V programs
79.1% Travel	70% Fashion
60.4% A Memorable day	70% A Memorable day
68% Transport	70% Religious Festivals
63.8% Gift giving	60% Transport
61.1% Celebrities	50% A Hotel Stay
<b>Topics STUDENTS did not feel comfortable talking about</b>	<b>Topics EXAMINERS felt students would not be comfortable talking about</b>
61.1% International News	70% Art
59.7% Plants and flowers	50% Concerts
55.5% Art	40% International News
52.7% Famous Buildings	40% Music
51.3% Concerts	20% Books

As can be seen above with the ranking of topics, there is both understanding and misunderstanding or convergence and divergence of perceptions.

Weddings and T.V programs are the undisputed perceived favourites for both , however there are some topics that rank in the top seven for students and not for examiners, such as Travel, Celebrities and Gift giving. The underlying assumption on the part of the examiners is possibly that the students would not be able to talk about travel because they do not have many travel opportunities; that they would not be able to talk about celebrities because this would not be in their sphere of experience not being exposed to western-style magazines and T.V programs and gift-giving because this also tends to be a western concept. Gifts of money are given to children at the two religious celebrations of Eid, but generally gifts such as birthday gifts and anniversary gifts are uncommon. On the other hand, surprisingly perhaps, the topic of Religious festivals, favoured by 70% of the examiners, is only thought to be an accessible topic by 31.9% of the

students. This could be because the students felt that they could not discuss any other festivals apart from their own or that they could not or would not want to discuss their own. Whatever the reason, the results are again evidence of the mismatch of the perceptions between both parties.

Turning to the topics that students either did feel comfortable talking about or would not want to talk at all about, the least favourite topic as perceived by the students is International news and events (61.1%) followed closely by Plants and Flowers (59.7%). In contrast, the perceived least favourite topic for students, according to the examiners, are two entirely different topics: Art (70%) and Concerts (50%). Whilst they agree that International news, Art and Concerts rank amongst the top five topics that students are least willing to talk about, there is disagreement about the other two topics in terms of rank order.

These topics listed above are typical IELTS-style topics for part two of the interview. The examiner has an element of choice as to which topic they choose from within the given binder of questions and it is assumed that if they choose (they can simply go through each topic in chronological order as all the topics are intentionally designed to offer equal access to all candidates) they will try to choose a topic with which they feel the candidate will be able to best showcase their language skills. If, however, as we have seen here, there is a mismatch between perceptions of accessible topics, the co-construction of language in the interview will be hampered, especially if the topic chosen is one of those ranked as the least accessible by the student. This then may have an impact on the grade outcome. The evidence from the video footage, in interview 4 for example, shows the examiner not only struggling with the fact that the candidate is veiled but also with the chosen topic of Art. The male examiner avoids looking up at the candidate and she is clearly struggling with the question: 'Tell me about Arabic forms of Art'. The candidate hesitates and utters a nervous laugh and is unable to continue until the examiner accommodates and gestures with his hand, air writing in the style of calligraphy. Between them, they co-

construct using tactics other than spoken language and eye contact or facial expression and arrive at conversation. It is not possible to say at this point whether the topic chosen in fact had an impact on grade outcome, but the evidence suggests it would. The same student would have to be interviewed again with a different topic (from the top ranking list) and the grades compared.

When applied on an international scale, the evidence suggests that there may be a need to address the issue of choice of topics. As Khan (2005) demonstrated in Bangladesh, and as I have demonstrated in the UAE, students' knowledge and schemata that they bring to the interview are dependent on their sociocultural background as are the questions that are created by the test makers dependent on their schema. The closer the perceptions of both are, the greater likelihood is it that the topics will be more accessible and not dependent on experiences or knowledge that the candidate has not shared.

Communication, as has already been argued, is a shared co-construction of meaning and for it to be valid, both parties, even in an assessment situation, should feel equally empowered to talk about the topic. This data correlates with Chen and Cegala's findings (1994) on topic management; for native speakers topics are developed as an information exchange based on shared meaning and understanding whereas for non-native speakers the struggle is not only with the linguistic element of communication but also with the need to find a shared conceptual common ground. This also supports the work of Nishida (1999) whose theory of cultural schemas suggests that intercultural competence is a cognitive process reliant on the development of schemas: the success of the communication between the candidate and interviewer in an EFL oral assessment could therefore be seen to be dependent on the degree of shared schemas between them. In short, interaction is more effective where interactants are from the same sociocultural and linguistic background and this balance is compromised in

EFL oral assessments because of the differences in schema brought to the interview.

As has been shown, it is very difficult to reach agreement on what is a generic topic and perhaps such an ideal is impossible to attain. Responses from the questionnaire seem to support the view that both students and examiners are aware that the cultural context of the assessment and each other's understanding of it, is part of the scaffolding underpinning the co-construction of meaning. Over two thirds of the respondents (69.4%) felt that it was important that the English teachers and examiner understood the cultural background of the students so that "she can respect our values", "know our culture and teach us equally with it" and "help us understand each other".

Similarly, 70% of the teacher examiner respondents thought it important not only to understand the culture but also to have a basic level of Arabic "to understand mistakes/errors students make, and the culture to understand which topics are better to use/not appropriate."; 'not so much the language, but understanding of the culture/tradition is very important for rapport'. If the examiners do have knowledge of Arabic language and culture: 'it reassures them'; 'again for rapport'; they like the 'respect' aspect of knowing about them/their culture/language.

From a teaching point of view, EFL resources and materials need to be culturally appropriate as they provide the bridge and supporting structure that allows accommodation of each party: the language from the teacher and the cultural schemata from the student. This follows through to the examination materials: language is not separable from culture, it is a social practice. Kramsch, (1993) notes that cultural awareness enables language proficiency, but here she is referring to the students. It seems to me that as the evidence has been presented, communication is the co-construction of meaning and therefore it is incumbent on both the 'co' parties, candidates



and the other (examiner, test creator, assessment boards) to foster cultural awareness.

The choice of topics, the data suggests, must be aligned with the cultural context and world knowledge of the test-takers. Similarly, the variety of language (English) tested should not be based on a native-speaker model. For most students, a native-like proficiency in the mode of standard British English or standard Australian English is unattainable and probably inappropriate (Kirkpatrick, 2007). Students should be tested on the variety of language that they are most likely to be using in their own cultural context. In the front cover of the IELTS research report, vol 7, IELTS is said to be 'for people who intend to study or work where English is the language of communication'. The implication seems to be that there is only one possible version of English. This sentiment was echoed by some of the examiners in the case study who felt that students should be aiming to speak standard British, American or Australian. However, as Kirkpatrick maintains, 'this makes any argument about the relative international intelligibility of such a model, 'frequently irrelevant' (p.191). The variety of language being tested, therefore, should reflect the sociocultural context and functional needs of its speakers.

### The impact of the veil

The veil, contrary to opinion expressed in the questionnaires by faculty and known to be current popular belief in Western Europe, at least as expressed by former leader of the British House of commons, Jack Straw, (2006) does not seem to hinder communication and therefore, in an oral communication assessment, does not seem to negatively affect the grade outcome. 40% of the faculty felt that the grade outcome would be worse where the student was veiled, 30% were unsure and 30% felt that there would be no impact on the grades. In the exam situation, in fact, wearing the veil had no impact at all on grade outcome. However, it must be noted that all the examiners had experience of working with veiled and partially-veiled students in the classroom context. This correlates with Coniam's findings (2005) that

students who wore face masks during the SARS outbreak, did not score less well in their oral assessments as a result. Feedback from the questionnaires also indicates that this is the view of the students who felt that wearing a veil would not result in a lower grade. Over half of the students, 68%, felt that wearing a veil would not affect the quality of communication during the assessment. This mismatch in perceptions regarding the veil and its impact on grade outcome is significant because it reflects the wider debate that is currently underway throughout Europe. Senator Francois-Noel Buffet (2010), for example, writing about the recent ban on the wearing of the full facial veil in France stated that:

‘To see without being seen undeniably constitutes a break in equality and at the same time an offence against human dignity, especially against women’s dignity’.

Buffet is expressing not only the concern about the veil being perhaps a symbol of oppression for women and a security risk, but also the general unease that appears to be felt that the veil, in hiding the face, also hinders communication. Opinion polls in Italy, Spain, Germany and Britain have indicated widespread public support for this sentiment. Whilst the scope of this small case study is not to consider the very contentious issue of the veil in terms of religious and political rights, I think the findings made in the study do indicate that there is an element of misunderstanding about those who wear the veil and its impact on their ability to communicate and be communicated with, alongside those who do not, with equal efficiency. In this case study at least, as far as EFL assessments are concerned, the students who wore a veil and those who did not, did not perceive the veil to be a hindrance to their communication skills and the grade outcome. And they were right. Given that the grade outcome is a measure of the success of the constructed communication between the examiner and the candidate, it follows that the argument against veils as being obstructive and hindering communication may need further investigation. In this case, the evidence points to the fact that wearing a veil has no significant effect on grade outcome in EFL oral assessments. However, in this context the participant

examiners in the research interviews were familiar with communicating with veiled students and this may have contributed to the lessening of any potential effects.

Following on from this, the issue of eye contact was yet another aspect of misplaced intercultural sensitivity. It is widely assumed that female Muslims are not comfortable with direct eye contact with males. In fact, all of the candidates used direct eye contact during the interviews. The data from the video footage confirms this. Data from the focus group also support this: when one student was asked whether the interviewer (male) allowed her to show off her speaking skills she replied:

Yeah. Because when they looked at er us and by eyes,eyes
contact,
and they smiled.
They make us feel comfortable when we talked.

The student is expressing her relief and the comfort she felt from the eye contact of her examiner. If the popular misconception about eye contact was followed through here and the examiner had averted or avoided direct gaze on the assumption that the candidate would feel uncomfortable, it would have had the opposite effect. Furthermore, significantly, as was also witnessed in the video data, it was the male examiner of a fully-veiled candidate who was more ill at ease with maintaining eye contact, and deliberately avoided it, despite concerted efforts by the candidate to establish eye contact. Here the lack of intercultural knowledge or misplaced sensitivity on the part of the examiner creates an unnecessary barrier for the candidate: in effect it is the examiner who is wearing the veil.

Gender of the examiner

Another interesting finding relates to the gender of the examiner. The results of grading differences show that same-gender markers (female examiners with female candidates) have the biggest discrepancy when considering the face-to-face and audio grading: an indicator to test creators that both males and females should rate an exam for maximum validity of the grading process. In four of the five interviews that involved mixed

gender grading, the results recorded by the female examiners were lower than those of the male examiners. In the data collected, female examiners are harsher markers than male and this matches the students' perceptions as indicated in the questionnaire responses. Although 65% recorded no particular preference for the gender of the examiner (65%), 27.7% said they would prefer a male examiner, as opposed to only 6.9% favouring a female examiner. Comments included 'the women are more strict' and 'I think that some females are biased'.

This is contrary to the prevailing Western image of a Muslim woman, veiled or non-veiled, in that the assumption would probably be that female candidates would prefer a female examiner. In fact the faculty questionnaire responses support this assumption too as 50% felt that students would prefer a female examiner compared with male (10%) and 40% were unsure or recorded no preference.

#### 4.4 Conclusions and implications for EFL testing and teaching

The Case Study was successful in that it presented evidence that suggests, in the interests of test validity, the need for regionally-specific English tests that can take into account accepted language variations and the local socio-cultural situation. Major stakeholders in EFL test design, test delivery and teaching and learning materials need to find ways for the international community to have an input in the process at a grass-roots level to ensure that validity, objectivity and therefore trust is maintained. The findings show that there is a mismatch in perceptions of appropriacy in terms of topic choices, for example.

It is no longer feasible to have international tests created and administered solely by Inner Circle stakeholders (Kachru 1990). There should be a consortium of contributors representing a broad range of the varieties of English, including input from test takers and examiners. Kirkpatrick (2007 p. 189) refers to this as a 'bilingual approach' one in which the yardstick by

which English competency is measured is in terms of L2 standards and not 'unattainable and inappropriate L1 standards'. Bamgbose (2001 p.359) reiterates this when he calls for 'communication across world Englishes to be seen in terms of accommodation between codes and in a multilingual context' to which I would add 'multicultural context'. The English taught, and tested, must be appropriate and relevant to the context in which it is being used and the needs of those in that context should not only drive the variety of English taught but also the generation of materials and resources suitable in that context. Spoken English cannot be seen simply in relation to its closeness to British or American or Australian English, for example; the reality is it is taught, tested and spoken by multilingual, multicultural non-native speakers and international assessment stakeholders must address this to retain test validity and reliability in the globalised context.

With regards to examiner training, despite being all being accredited examiners there were grade discrepancies. In Nakatsuhara's study, (2008) the IELTS trained examiner who followed the test rubrics had a candidate who scored less well and Nakatsuhara concludes that the trained examiner did not provide the opportunity for the best performance from the candidate'. This has implications for examiner training; examiners must be made aware of how to create the best opportunities for candidates to showcase their language skill, but this is also a responsibility of test writers to facilitate this through informed and culture-sensitive writing of test rubrics.

Also, as has been seen, there are often grade discrepancies even between trained, accredited examiners and it could be argued that a grade allocation by one individual alone is not a valid assessment. The findings reveal that the least discrepancy occurred where there was both a male and female examiner and Mullen (1978) also concluded that two assessors result in a more reliable assessment, although he does not reveal whether the gender is important or not. This has implications for both the examination process and teaching and learning strategies. If there were two assessors in the face

to face exam, one could focus on the interview rubrics, the other in observing and grading the whole interaction, not simply the linguistic exchange.

Moreover, teachers need to be made aware of the significant impact that paralinguistic communication has on the grade outcome and perhaps adapt their teaching of speaking skills accordingly. In the audio recordings, the examiner is merely an observer. The face to face interview is not only a dynamic exchange of speech but also of cultures, past experiences, expectations and aspirations as communicated through body language, facial expression and physical interaction with the environment and each other. The question therefore needs to be asked: is it realistic to be using the same grading criteria in both scenarios? If not, can the grading criteria or grading process be adapted accordingly? Should there be a separate set of criteria for audio recordings? This is an area that would benefit from further research as effectively the two methods of assessment are assessing two different modes of communication.

One of the areas of grade discrepancy was in the category of grammatical range and accuracy. All three of the differences in grades between GR occurred as higher grades for the face to face interview than the audio grading. Interestingly, all three ranges of veiling were evident here: interview four (veiled, eyes hidden); interview six (veiled eyes visible); interview seven (non-veiled) and with two different interviewer gender ratios: four & six M/F; seven F/F. This would suggest that GR is probably the least affected of the grading categories in terms of gender and veiling but, surprisingly, all three were a band lower when audio graded. As with LR, it might be assumed that GR is more finite and tangible, countable even and so there would be no added communicative value assignable to a face to face interview. With this limited data of three interviews, it is impossible to make generalizations but this is an area that would benefit from further research.

Byrne (cited in Giles and Smith, 1979) suggested that the closer our own ideals in terms of attitudes and beliefs are, the more likely we are to converge with the person to whom we are speaking. That said, in EFL assessments, the closer the examiner and candidate are in attitudes and beliefs, the less linguistic and behavioural effort needs to be utilized to ensure comprehension. Yoshida (2001) suggests that pronunciation, speech rate and message content are three possible levels of 'intralingual' convergence and depending upon the negotiated balance of these aspects of speech, understanding will take place to a greater or lesser extent. As discussed earlier in the literature review, the danger of attempting to converge when the recipient is not receptive, is that it may be viewed as 'patronising, condescending or even threatening' (Yoshida 2001, p36). This is of critical importance for both candidate and examiner: the candidate will want to maximize their grade by doing whatever they perceive that the examiner needs to see and hear and the examiner's job is to allow for the opportunity for the candidate to offer evidence of their linguistic competence but if either is considered to be over-accommodating and converging, this can negatively affect the interview and hence potentially the grade.

It is therefore important that further research is carried out in this field to help inform the test creators and thereafter the training of both examiners and candidates. Examiners need to understand the discourse of accommodation as this will further add to the validity of the test, particularly in the case of IELTS which is a global testing system, and ensure more consistency of marking. Further, candidates also need to be aware of the discourse of accommodation to ensure that they can maximize the interview test opportunity. This has implications, not only for stakeholders in international EFL assessments but also for the publishers of teaching materials so that teachers can be empowered with appropriate resources.

Training of EFL examiners and teachers needs to include strategies for identifying situations where students are not achieving success, either in a test or a class situation, not because of a lack in linguistic skill but because of a possible imbalance of schemata that is affecting their ability to accommodate and converge with the teacher or assessor. At the same time, students need to be knowledgeable of, and equipped with, the non-linguistic skills underpinning speech to facilitate the co-construction of meaning. Maintaining a level of eye contact, for example, was one of the areas where misplaced sensitivity caused divergence in communication strategies when the veiled student (interview four) strove to gain the attention of the examiner but, since she was a veiled student, the examiner rarely looked directly at her on the assumption that a) she did not want to be looked at for religious reasons or b) she could not see him from behind the veil. Both assumptions were false. The student, in anecdotal evidence, revealed that this experience was extremely frustrating for her and she felt that she must have been boring the examiner or that he was not paying attention to her. The examiner, again anecdotally, revealed that he had felt very uncomfortable with the situation and was keen not to offend or upset the student in any way. In this case, the nonlinguistic behavior impaired the co-construction of meaning and could possibly have been avoided if due thought had been given to the cultural context of the test. Perhaps part of the introductory rubric might include instructions for how to deal with veiled/partially veiled students in relation to eye contact.

In the teaching context, when teachers are preparing students for the oral assessment, it is important that non-linguistic behaviour be discussed, not only that of the students but also that of the examiners. What should they expect from the examiner and what is the most appropriate way to respond?

In summary, the findings challenge several commonly held assumptions about these female, Muslim, Arab students at the centre of the case study and the context of EFL learning and international assessment. Stereotypes have also been challenged and broken down.



So, to answer the title question, ‘Emirati females and an international EFL oral proficiency test, does one size really fit all?’ , the answer must be no. EFL oral proficiency tests must evolve in the light of this research. The findings have applicability and a generic value to all EFL assessment contexts in cultures that are dissimilar to the culture embedded in the assessment materials. If English is to continue to be the language of globalization, it will continue, increasingly, be taught and tested in cultural contexts that are currently alien to the context of the origin of the assessment tools. Kirkpatrick, (2007, p. 197) recognizes the need for multilingual and multicultural EFL teachers to reflect the fact that ‘variation is natural, normal and continuous.’ This level of variation must now follow through into EFL assessment and include the discourse of accommodation. One size does not appear to ‘fit all’.

#### 4.5 Opportunities for further research

The challenge for researchers, based on the evidence from the case study, is now to further refine theories about oral communication, to apply and develop CAT as a tool for understanding more about interaction generally and intercultural interaction specifically. Further independent research is needed into international EFL oral tests, such as IELTS (already committed to their own highly-respected research) to inform and help develop the assessment process in the light of the increasing demand, globally, for certification in English. Consumers (candidates and test users) need to feel confident that the test they are taking is valid, reputable and internationally recognized but that also it is locally relevant to their needs and the variety of English that they will be using.

Awareness-raising of preconceptions about cultural differences by inviting greater input from test users around the globe would help inform the choices test creators make when compiling test materials with particular emphasis on topic appropriacy. It is also important to note that no culture is static and that, as Nishida (1999) points out, cultural schemas evolve according to the experiences and interactions of the individual. The field of

intercultural communication is therefore like a shifting sand to try to research; when identical messages can be interpreted differently depending upon your cultural schemas there is, in effect, an infinite number of possible interpretations. Topics in assessments and in teaching materials must be both inclusive and culturally appropriate, the challenge being here is not to sanitise the resources so much that they become bland and as equally difficult to discuss as a culturally inappropriate topic might be (Gray, 2005).

The goal for EFL test writers is to create valid assessments that test all competencies accurately and allow all candidates to achieve grade outcomes that reflect their ability to communicate in English. To this end further research needs to be conducted to help shed light on defining the techniques and input assessors should employ, to guide their discourse of accommodation. They need to be made aware of the impact they have on the co-construction of the speech act and the responsibility that accompanies that role in being able to award a valid and reliable speaking grade to the candidate. Further research is therefore needed too into the candidate's discourse accommodation, for the same reasons.

More research also needs to be conducted into the varieties of English and how this impacts on teaching and learning and international testing. As seen in the findings of the case study, some examiners felt that students should speak a 'standard' form of British or American English. This has implications for the grading criteria of international EFL assessments and the training of examiners, in particular for categories such as pronunciation and accuracy and appropriacy. Who should be the arbitrator of appropriacy and according to which standard?

CAT, as a tool for exploring communication discourse and interactional competence, needs to be further developed and the measurement yardsticks more clearly defined to increase its scope and applicability. It would be particularly interesting to see it applied to online or virtual testing.

Ironically, virtual testing is a reality and will no doubt generate other issues

concerning accommodation strategies and the co-construction of meaning that are both similar to and different from those experienced in face-to face contexts. Virtual tests where faces are visible, for example via webcams and tests where only an audio recording is accessible, may present the same issues as those found in the case study research on face-to face interviews. This provides a whole new area of research. Develotte, Guichon and Vincent (2010) point to the fact that teachers will need to be trained in both pedagogical and online communication strategies and that the non-verbal strategies that we use in face to face communication may be different in online teaching or testing environments. Further, they suggest that perhaps using webcams as a mode of communication will lead to “the development of a specific interactional body language, one that is adapted to this realm of perception” (Develotte et al, 2010, p.5). Oral testing via a webcam would present a further dimension to the context. As Lamy and Hempel (2007) point out, webcam images can be poor and only usually show the head and shoulders of the person on camera thereby reducing the contextual information available such as physical environment, whole body language and clothes. Furthermore, the connection may be such that there is an echo or slight delay in the audio thus impairing the simultaneous nature of the communication. In fact, Develotte et al’s research (2010) found that teacher trainees chose to stop using the webcam images available to them on their course, using them as a possible resource rather than a preferred form of communication (Develotte 2010 p. 16). It seems that out of choice, people prefer to communicate face to face and that an EFL oral assessment, mediated online via an avatar or webcam image, would require careful training for test writers, examiners, raters and test takers in acknowledgement of the different set of communication strategies required in this context.

The way forward, as Luoma (2004 p. 191) concludes is through ‘action, reflection and reporting’. It is important to pilot international EFL speaking tests in a variety of socio-cultural contexts (and online environments) and with multicultural examiners, to examine not only the applicability of the

test materials as indicators of linguistic ability but also their applicability in relation to the learning context. We need to learn from the experiences of both examiners and candidates about their reactions to the test materials and also to each other.

Further studies could look at the co-construction of meaning through discourse and behavioural analysis of 'examiner talk' and 'candidate talk' to try to identify which features, if any, are indicators of success in terms of grade outcome. Brown and Hill (2007) identify the challenge for IELTS, in particular, as the need 'to decide what behavior is appropriate and to ensure that it occurs' (p.56). Examiners in any oral assessment must be trained to be self-aware as 'interviewer talk is not neutral' (p.56). The same applies to candidate behaviour, candidates need to be more self-aware so that they can converge more easily with their assessor.

Communication is a multifaceted, multi-layered complex phenomenon: to assess speech in isolation, as currently happens in EFL oral assessments, is relegating the act of communication to a mere exchange of sounds and punctuated rhythms. There is little consideration for the discourses of accommodation or the negotiation and co-construction of meaning which gives depth and validity to the spoken word. The two are inseparable and interdependent, and whilst troublesome perhaps to tackle, must be addressed if oral testing is to be an assessment of communicative ability and not simply linguistic skill. It follows that further research needs to be done in the differing nature of assessing face to face communication compared with audio recordings. Currently, examiner and candidate monitoring is conducted, after the face to face event, by listening to an audio recording. Grading criteria needs to be developed that reflects the fact that the face to face interaction is a complex negotiation of meaning both linguistically and schematically.

To sum up, as Kramsch (2006) so aptly writes:

.. it is no longer appropriate to give students a tourist-like competence to exchange information with native speakers of national languages within well-defined national cultures. They need a much more sophisticated competence in the manipulation of symbolic systems.....Language learners are not just communicators and problem solvers, but whole persons with hearts, bodies, and minds, with memories, fantasies, loyalties, identities (Kramsch 2006, p251).

Here Kramsch is highlighting the multidimensional aspect of communication. Language learners, teachers, assessors and test writers perhaps need to rethink the concept of linguistic skill and see language competency as including all these aspects. Furthermore, greater attention needs to be focused on the purpose of the test itself: a test of global English must also represent the needs of the local context in which it is being administered, as the Case Study findings have presented. In the future, more international tests of English may be conducted online and this will present other issues to do with context and mode, as already discussed. To conclude, without a valid and reliable test process applicable or adaptable to a variety of contexts, the test results do not mean much. The goal must therefore be: a global test for local needs.



## **IMAGING SERVICES NORTH**

Boston Spa, Wetherby  
West Yorkshire, LS23 7BQ  
[www.bl.uk](http://www.bl.uk)

# BLANK PAGE IN ORIGINAL

## REFERENCES

- Aitchison, J. (ed.) (2001) *Language Change: Progress or Decay?*, 3rd edition. Cambridge, Cambridge University Press.
- Alexander, R. (2000) *Culture and Pedagogy*. USA, Blackwell Publishers Ltd.
- Alptekin, C. (1993) 'Target language in EFL materials', *ELT Journal* 47/2:136-43.
- Alred, G and Byram, M. and Fleming, M. (ed) (2006) *Education for Intercultural Citizenship*. Clevedon, England, Multilingual Matters Ltd.
- Altbach, P. G. (2008) 'The Imperial Tongue: English as the Dominating Academic Language', *International Educator*, 17(5) p.56 - 59.
- Arnold, J. (ed) (1999) *Affect in Language Learning*. Cambridge, UK, Cambridge University Press.
- Ayoko, O and Hartel, C. and Callan, V. (2002) 'Resolving the puzzle of productive and destructive conflict in culturally heterogeneous workgroups: A communication accommodation theory approach'. *International Journal of Conflict Management*, vol.13, no.2, p.165.
- Bachman, L. (1990) *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Bachman, L. (2004) *Statistical Analyses for Language Assessment*. Cambridge, UK, Cambridge University Press.
- Badry, F. (2007) *Research in ELT Contexts*. Dubai, UAE, TESOL Arabia.
- Bamgbose, A. (2001) World Englishes and globalization. *World Englishes*, Vol. 20 (No.3), p.357-364.
- Barber, C. (2000) *The English Language: A Historical Introduction* (2nd edition). Cambridge, Cambridge University Press.
- Baxter, L and Braithwaite, D. (ed) (2008) *Engaging Theories in Interpersonal Communication*. Thousand Oaks, California, Sage Publications Inc.
- Bejar, I (2010) *Can Speech Technology Improve Assessment and Learning? New capabilities May Facilitate Assessment Innovations*. ETS Research Report. Vol 25. August 2010. Princeton, NJ, Educational Testing Service.
- Bell, J. (2009) *Doing your Research Project*. England, Open University Press.

- Bell, R (2006) Paper 057, 'Teachers' assessment practices for classroom diversity' IAEA 2006 32nd International Association for Educational Assessment (IAEA) [http://www.iaea.info/documents/paper\\_1162a1ffcd.pdf](http://www.iaea.info/documents/paper_1162a1ffcd.pdf)
- Berwick, R and Ross, S (1996) Cross-cultural pragmatics in oral proficiency interview strategies. In Milanovic, M and Saville, N (eds) *Performance Testing, Cognition and Assessment, Studies in Language Testing 3, Selected Papers from the 15<sup>th</sup> Language Testing Research Colloquium, Cambridge and Arnheim*. Cambridge: UCLES/Cambridge University Press, 34-54.
- Bex, T. and Watts, R. (ed) (1999) *Standard English: the Widening Debate*. London, Routledge.
- Block, D. and Cameron, D. (ed) (2002) *Globalization and Language Teaching*. Abingdon, Oxon, Routledge
- Bristol-Rhys, J. (2009) Emirati Historical Narratives. *History and Anthropology*, Vol.20, No2, June 2009, p.107-121.
- Broughton, G et al (1980) *Teaching English as a Foreign Language: Second Edition*. London, Routledge & Kegan Paul Ltd.
- Brown, A. and Hill, K. (2007a) 'Interviewer style and candidate performance in the IELTS oral interview', in Taylor, L. and Falvey, P. (eds), *Studies in Language Testing 19*, p.37-61. Cambridge, Cambridge University Press.
- Brown, A. (2007b) 'An investigation of the rating process in the IELTS oral interview', in Taylor, L. and Falvey, P. (eds), *Studies in Language Testing 19*, p. 98-138. Cambridge, Cambridge University Press.
- Buffet, F (2010) 'Banning the burqa: France votes against the veil' Euronews 15.09.10 <http://www.euronews.net>.
- Bugelski, B. (1956) *The Psychology of Learning*. New York, Holt, Rinehart and Winston, Inc.
- Bull. P (2001) 'State of the art: Nonverbal communication'. *The Psychologist*. Vol 14, No 12, 2001, 644 – 647.
- Burgess, H and Siemenski, S and Arthur, L. (2006) *Achieving Your Doctorate in Education*. London, Sage Publications.
- Buttjes, D. and Byram, M. (ed) (1991) *Mediating Languages and Cultures: Towards an Intercultural Theory of Foreign Language Education*. Clevedon, England, Multilingual Matters Ltd.



- Byram, A. (2001) *Social Research Methods*, Oxford, Oxford University Press.
- Byram, M. (1989) *Cultural Studies in Foreign Language Education*. Clevedon, England, Multilingual Matters Ltd.
- Byram, M. (1997) *Teaching and Assessing Intercultural Communicative Competence*. Clevedon, England, Multilingual Matters Ltd.
- Byram, M. (2008) *From Foreign Language Education to Education for Intercultural Citizenship: Essays and Reflections*. Clevedon, England, Multilingual Matters Ltd.
- Byram, M. and Nichols, A. and Stevens, D. (ed) (2001) *Developing Intercultural Competence in Practice*. Clevedon, England, Multilingual Matters Ltd.
- Cameron, D (2000) *Good to Talk? Living and Working in a Communication Culture*. London, Sage.
- Canagarajah, A.S (1999) *Interrogating the Native Speaker Fallacy*. Baine, G (ed) *Non-Native Educators in English Language Teaching*. New Jersey, Lawrence.
- Carrell P, and J. Devine, and D. Eskey (1988) *Interactive Approaches to Second Language Reading*. Cambridge, Cambridge University Press.
- Carroll, D. (1986) *Psychology of Language*. California, Brooks/Cole Publishing Company (Wadsworth, Inc.).
- Charlesworth, Z (2008) Learning styles across cultures: Suggestions for educators. *Education and Training*, 50 (2), 115-127.
- Chen, L. and Gegala, D. (1994) 'Topic Management, Shared Knowledge and Accommodation. A Study of Communication Adaptability.' *Research on language and Social Interaction*, vol. 27, no.4, p. 389 – 417.
- Cheng, X (2000) Asian students' reticence revisited. *System*, 28, 435-446.
- Chew, P (1999) 'Linguistic Imperialism, globalism, and the English language'. *English in a changing World*. AILA Review 13, p.37-47.
- Chomsky, N. (2006) *Language and Mind*. New York, Cambridge University Press.
- Chomsky, N. (2007) *On Language*. New York, The New Press.
- Chomsky, N. (2008) *The Essential Chomsky*. Great Britain, The Bodley Head.

- Clarke, M. (2006) 'Beyond antagonism? The Discursive Construction of "New" Teachers in the United Arab Emirates'. *Teaching Education*, Vol.17, No.3, September 2006, p.225-237.
- Clarke, M. (2008) *Language Teacher Identities*. Clevedon, Multilingual Matters.
- Coates, J. (1993) *Women, Men and Language*. London, Longman.
- Cohen, L. and Manion, L. and Morrison, K. (2000) *Research Methods in Education: (Fifth Edition)*. London, Routledge Falmer.
- Coniam, D. (2005) 'The impact of wearing a face mask in a high-stakes oral examination: an exploratory post-SARS study in Hong Kong'. *Language Assessment Quarterly*, 2 (4) 2005.
- Coombe, C. and Davidson, P. and Lloyd, D. (ed) (2003) *Proceedings of the 5th & 6th Current Trends in English Language testing (CTELT) Conferences (Vol. 3)*. Dubai, UAE, TESOL Arabia.
- Corbett, J. (2003) *An Intercultural Approach to English Language Teaching*. Clevedon, England, Multilingual Matters Ltd.
- Crystal, D. (2003) *English as a Global Language: Second Edition*. Cambridge, UK, Cambridge University Press.
- Cupach, W. and Imahori, T. (2005) 'Identity Management Theory: Facework in Intercultural Relationships', in Gudykunst, W (ed) *Theorising about Intercultural Communication*, p.195-210, U.S.A, Sage Publications.
- Cushner, K. and Brislin, R. (1986) *Intercultural Interactions: A Practical Guide (Second Edition)*. Thousand Oaks, California, Sage Publications Inc.
- Dahan, L. (2007) 'English as an International language in the Arabian Gulf: Student and Teacher Views on the role of culture', in Midraj, S. and Jendli, A. and Sellemi, A. (eds) *Research in ELT contexts (2007)*, Dubai, TESOL Arabia publications.
- Davidson, C. (2005) *The United Arab Emirates, a Study in Survival*. London, L. Rienner.
- Davidson, P et al (ed) (2007) *Teaching and Learning Vocabulary in Another Language*. Dubai, UAE, TESOL Arabia.
- Davies, A. (2003) *The Native Speaker: Myth and Reality*. Clevedon, England, Multilingual Matters Ltd.
- Davies, A. (2008) *Assessing Academic English: Testing English Proficiency 1950-1989- the IELTS Solution*. Cambridge, UK, Cambridge University Press.

- Davies, A. and Elder, C. (ed) (2004) *The Handbook of Applied Linguistics*. USA, Blackwell Publishers Ltd.
- Davies, A et al. (1999) *Dictionary of Language Testing*. Cambridge, UK, Cambridge University Press.
- Demont-Heinrich, C. (2009) 'Language, Globalisation, and the Triumph of Popular Demand: The Discourse of Populism in American Prestige Press coverage of the Global Hegemony of English'. *The Communication Review*. 12:p.20-49.
- Denzin, N.K. (1969) 'Symbolic interactionism and ethnomethodology: A proposed synthesis'. *American Sociological Review*, 34, p.922-934.
- Develotte, C and Guichon, N and Vincent, C (2010) 'The use of the webcam for teaching a foreign language in a desktop videoconferencing environment'. *Journal RECALL*, Vol. 22, Issue 3. CUP.
- Donskoi, V. (2006) 'Systemic theory of language competition'. *Journal of Language and Politics* 5:2 (2006), p.277-298.
- Dulay, H. and Burt, M. and Krashen, S. (1982) *Language Two*. Oxford, Oxford University Press.
- Edge, J. (1994) 'ELT culture and communication' *IATEFL newsletter*. 126, p.18-19.
- Edwards, S. and Zahnd, E. and Willis, G. and Grant, D. and Lordi, N. and Fry, S. (2004) 'Behaviour coding across multiple languages: The 2003 California Health Interview survey as a case study', *Conference Papers – American Association for Public Opinion Research*, 2004 Annual Meeting, Phoenix, AZ.
- Ehrman, M. (1996) *Second Language Learning Difficulties: Looking Beneath the Surface*, Thousand Oaks, C.A, Sage Publications.
- Elder, C et al. (ed) (2001) *Experimenting with Uncertainty (Essays in Honour of Alan Davies)*. Cambridge, UK, Cambridge University Press.
- Elder, C. and Wigglesworth, G. (2003) 'An investigation of the effectiveness and validity of planning time in Part 2 of the IELTS speaking test', in *IELTS Research Reports*, Volume 6, p. 13-40, 2006, British Council/IELTS Australia Pty Limited.
- Ellingsworth, H (1983) 'Adaptive intercultural communication', in Gudykunst, W. (ed) *Intercultural communication theory*, p195-204. Beverly Hills, CA, Sage.

Ellis, R (1990) *Instructed Second Language Acquisition*. Oxford: Blackwell

Faerch, C. and Kasper, G. (1984) 'Two ways of defining communication strategies'. *Language Learning*, vol.34, p.45 – 63.

Fairclough, N. (1989) *Language and Power*. London, Longman.

Fairclough, N. (2006) *Language and Globalization*. Abingdon, Oxon, Routledge.

Feldman, R. and Rime, B. (ed) (1991) *Fundamentals of Nonverbal Behavior*. Cambridge, UK, Cambridge University Press & Maison des Sciences de l'Homme.

Feldman, R. and Philippot, P. and Custrini, R. (1991) 'Social competence and nonverbal behavior'. In R.S. Feldman and B. Rimé (Eds.), *Fundamentals of Nonverbal Behavior* (p.329-350). Cambridge, Cambridge University Press.

Fromkin, V. and Rodman, R. (1993) *An Introduction to Language: Fifth Edition*. New York, Holt, Rinehart and Winston, Inc.

Fulcher, G. (1996) 'Does thick description lead to smart test? A data-based approach to rating scale construction'. *Language Testing*, vol. 13, p.208 – 38.

Fulcher, G. (2003) *Testing Second Language Speaking*. Great Britain, Pearson Education Limited.

Fulcher, G. (2010) *Practical Language Testing*. Great Britain, Hodder Education.

Fulcher, G. and Davidson, F. (2007) *Language Testing and Assessment*. Abingdon, Oxon, Routledge.

Fulcher, G. and Marquez, R. (2003) 'Task difficulty in speaking tests'. *Language Testing*, vol.20. no.3, p.321 – 344.

Gallois, C. and Ogay, T. and Giles, H. (2005) *Communication Accommodation Theory*, in Gudykunst, W (ed), *Theorising about Intercultural Communication*, U.S.A, Sage Publications.

Gass, S. and Lefkowitz, N. (1995) *Varieties of English*. USA, The University of Michigan Press.

Gibbon, P. (1993) *Learning to learn in a Second language*. London, Heinemann.

Giglioli, P. (ed) (1972) *Language and Social Context*. England, Penguin Books.

- Giles, H. (2008) 'Accommodating Translation Research', *Journal of Applied Communication Research*, vol. 36, no. 2, p.121 – 127.
- Giles, H. and Smith, P. (1979) 'Accommodation theory: Optimal levels of convergence'. In H. Giles., R. St Clair (Eds), *Language and social psychology* (p. 45-65) Baltimore, University Park Press.
- Goffman, E. (1981) *Forms of Talk*. Philadelphia, University of Pennsylvania Press.
- Goodwin, M (1990) *He said she said; talk as social organization among Black children*. Bloomington, Indiana University Press.
- Graddol, D. and Cheshire, J. and Swann, J. (1987) *Describing Language*. England, The Open University Press.
- Gray, J (2005) 'The global coursebook in English Language Teaching' in Block, D and Cameron, D, *Globalisation and Language Teaching*, Oxon, Routledge.
- Gregory, S. and Webster, S. (1996) 'A Nonverbal signal in voices of interview partners effectively predicts Communication Accommodation and Social Status perceptions', *Journal of Personality and Social Psychology*, vol. 70, p.1231 – 1240.
- Gudykunst, W. (ed) (2005) *Theorizing About Intercultural Communication*. Thousand Oaks, California, Sage Publications Inc.
- Gunzenhauser, G. W. (1996). *Assuring an appreciation for student diversity: alternative to teacher education field experiences*. [Electronic version]. US: Illionis.
- Hall, E. (1959) *The Silent Language*. New York, Anchor Books, A Division of Random House, Inc.
- Hall, E. (1976) *Beyond Culture*. New York, Anchor Books, A Division of Random House, Inc.
- Halliday, M.A.K. (1978) *Language as a Social Semiotic*. London, Edward Arnold.
- Halliday, M.A.K. (1993) 'The act of meaning' in Atlis, J. (ed), *language, communication and social meaning* (p.7-21). Georgetown University Roundtable on Language and Linguistics, 1992. Washington D.C, Georgetown University Press.
- Halliday, M. (2003) *On Languages and Linguistics*. London, Continuum International Publishing Group

- Hamers, J. and Blanc, M. (1989) *Bilinguality and Bilingualism: Second Edition*. Cambridge, UK, Cambridge University Press.
- Hampel, R. and Hauck, M. (2006) 'Computer-mediated language learning: making meaning in multimodal virtual learning spaces'. *The JALT CALL Journal*, 2006, vol.2 No.2, p.3-18.
- Harindranath, R. (2006) *Perspectives on Global Cultures*. England, Open University Press.
- Hawkey, R. (2006) *Impact Theory and Practice (Studies of the IELTS test & Progetto Lingue 2000)*. Cambridge, UK, Cambridge University Press.
- Hawley, D (1970) *The Trucial States*. London, Allen and Unwin.
- Heffernan, G. (2008) 'An Investigation of Diachronic Change in Communication Accommodation', *Journal of Language and Social Psychology*, vol. 27, no.1, p. 86-93.
- Hinkel, E. (ed) (1999) *Culture in Second Language Teaching and Learning*. Cambridge, UK, Cambridge University Press.
- Hofstede, G. (1984) *Culture's consequences: International differences in work-related values*. Newbury Park, CA, Sage.
- Holliday, A. (2002) *Doing and Writing: Qualitative Research*. London, Sage Publications Ltd.
- Holliday, A. and Hyde, M. and Kullman, J. (2004) *Inter-cultural Communication: an Advanced Resource Book*. Abingdon, Oxon, Routledge.
- Hornsey, M and Gallois, C (1998) 'The impact of interpersonal and intergroup communication accommodation on perceptions of Chinese students in Australia', *Journal of Language and Social Psychology*, 17, p. 323-347.
- Horvitz, E. and Horvitz, M. and Cope, J. (1986) 'Foreign language classroom anxiety', *Modern Language Journal*, 70, p.125-132.
- Hughes, A. (1989) *Testing for Language Teachers*. Cambridge, UK, Cambridge University Press.
- Imahori, T. and Lanigan, M. (1989) 'Relational model of intercultural communication competence', *Intercultural Journal of Intercultural Relations*, 13, p. 269-286.
- Ingram, D. and Bayliss, A. (2004) 'IELTS as a predictor of academic language performance, part 1', *IELTS Research Reports*, Volume 7, 2007, p. 137-204. U.K, British Council, Australia, IELTS Australia Pty Limited.

- Issit, S. (2008) 'Improving scores on the IELTS speaking test', *ELT Journal*, vol.62. no.2, p.131 – 138.
- Jenkins, S. and Parra, I (2003) 'Multiple Layers of Meaning in an Oral Proficiency Test: The Complementary Roles of Nonverbal, Paralinguistic, and Verbal Behaviors in Assessment Decisions', *The Modern Language Journal*, 2003, 87, i p.90-107.
- Johnson, K. and Morrow, K. (ed) (1981) *Communication in the Classroom*. England, Longman Group Ltd.
- Jones, J. (2004) 'A Cultural Context for Language Anxiety', *EA Journal*. Volume 21, no.2., p. 30-39.
- Kachru, B. (1990) *The Alchemy of English*. Chicago, University of Illinois Press.
- Kachru, B (1992) *The Other Tongue: English across cultures*. Oxford, Oxford University Press.
- Kachru, B. and Smith, L. (ed) (2001) *World Englishes (Vol.20)*. USA, Blackwell Publishers Ltd.
- Kachru, B. and Smith, L. (ed) (2002) *World Englishes (Vol.21)*. USA, Blackwell Publishers Ltd.
- Kayman, M. (2004) 'The state of English as a global language: communicating culture', *Textual Practice*. 18 (1) 2004, p.1-22.
- Kazim, A. (2000) *The United Arab Emirates A.D 600 to the present: A sociodiscursive Transformation in the Arabian Gulf*. Dubai, Gulf Book Center.
- Kelly, J.B. (1968) *Britain and the Persian Gulf 1795-1880*. London: Clarendon press.
- Khallaf, S. (2002) 'Globalisation and heritage revival in the Gulf: an anthropological look at Dubai Heritage Village', *Journal of Social Affairs*, Vol 19, no 75 p.13-42.
- Khallaf, S. (2009) 'Teaching Science and Maths in English hurts Arabic, schools warn.'  
<http://1426.blogspot.com/2009/04/teaching-science-maths-in-english-hurst.html>
- Khan, R (2005) 'The IELTS Speaking Test: Analysing Cultural Bias', *IATEFL Testing*, October, 2005 p.11-19.
- Khoddami, F. (2006) 'Iranian Teachers' Attitude towards the Intercultural Dimension of TESOL', unpublished paper presented at CTELT (Tesol

Arabia) conference, 2006, Dubai.

Kirkpartick, A. (2007) *World Englishes: Implications for International Communication and English Language Teaching*. Cambridge, Cambridge University Press.

Kramsch, C. (1986) 'From language proficiency to interactional competence?' *Modern Language Journal*, 70.4 p.366-371.

Kramsch, C. (1993) *Context and Culture in Language Teaching*. Oxford, Oxford University Press.

Kramsch, C. (1998) *Language and Culture*. Oxford, Oxford University Press.

Kramsch, C and Thorne, S (2002) 'Foreign language learning as global communicative practice' p83 – 100, in Block, D and Cameron, D (eds) *Globalization and language teaching*, London, Routledge.

Kramsch, C. (2006) 'From Communicative Competence to Symbolic Competence'. *Modern Language Journal*. 90.2, p.249-252.

Kress, G. and Van Leeuwen, T. (2001) *Multimodal Discourse*. London, Hodder Education.

Kreuger, R. (1994) *Focus group interviews: A practical guide for applied research*. 2d ed. Thousand Oaks, CA: Sage.

Kress, G. and Van Leeuwen, T. (2001) *Multimodal Discourse: the Modes and media of Contemporary Communication*. London, Edward Arnold.

Kubota, R. and Ward, L. (2000). Exploring Linguistic Diversity through World Englishes. *The English Journal*, vol. 89 No. 6, p.80 - 86.

Lamy, M and Hampel, R (2007) *Online Communication in Language Learning and Teaching*. Basingstoke: Palgrave Macmillan.

Lange, D. and Paige, M. (ed) (2003) *Culture as the Core: Perspectives on Culture in Second Language Learning*. USA, Information Age Publishing Inc.

Lankshear, C. and Knobel, M. (2004) *A Handbook for Teacher Research*. Berkshire, England, Open University Press.

Lanteigne, B (2006) 'Regionally Specific Tasks of Non-Western English Language Use' *TESL-EJ*, September, 2006, Volume 10, Number 2.

Lanteigne, B (2007) 'A different culture or just plain rude?' *English Teaching: Practice and Critique* September, 2007, Volume 6, Number 2



Lantolf, J. (2000) 'Introducing sociocultural theory', in Lantolf, J. (ed) *Sociocultural Theory and Second Language Learning*, p. 1-26, Oxford, OUP.

Lantolf, J. and Pavlenko, A. (1998) '(S)econd (L)anguage (A)ctivity theory: understanding second language learners as people.' In M. Breen (ed.), *Learner Contributions to Language Learning: New directions in research*. Harlow, Essex: Pearson Education Limited, pp.141-158.

Lazaraton, A. (1996) 'Interlocutor support in oral proficiency interviews: the case of CASE', *Language Testing*, vol.13, no.2, p.151.

Lazaraton, A. (2006) 'Process and outcome in paired oral assessment', *English Language Teachers Journal*, vol.60, no.3, p.287 – 289.

Lightbrown, P. and Spada, N. (1999) *How Languages Are Learned: Second Edition*. Oxford, Oxford University Press.

Lindlof, T. and Taylor, B. (2002) *Qualitative Communication Research Methods*. London, Sage Publications Inc.

Lock, S. 'A new generation for a new world'. *The National*, October 12, 2008.

Long, M and Crookes, G (1992) 'Three approaches to task-based syllabus design'. *TESOL Quarterly* 26/1:27-47.

Lowenberg, P. (2002) 'Assessing English proficiency in the Expanding Circle'. *World Englishes*, vol.21, Nov3, p.431-435.

Lumley, T and Brown, A (1996) 'Specific-purpose language performance tests: task and interaction', in Wigglesworth, G and Elder, C (eds) *Australian review of Applied Linguistics Series S*.

Lumley, T. and O'Sullivan, B. (2005) 'The effect of test-taker gender, audience and topic on task performance in tape-mediated assessment of speaking', *Language Testing*, vol.22, no.4, p.415 – 437.

Luoma, S. (2004) *Assessing Speaking*. Cambridge, UK, Cambridge University Press

Lustig, M. and Koester, J. (2006) *Intercultural Competence: Interpersonal Communication Across Cultures*. USA, Pearson Education, Inc.

MacAndrew, R. (2002) *English Observed*. Cambridge, Boston, USA, Heinle, Thomson Corporation.

- Madsen, H. (1983) *Techniques in Testing*. Oxford, Oxford University Press.
- Master, P. (1998). 'Positive and Negative Aspects of the Dominance of English', *TESOL Quarterly*, 32, 4, Winter 1998, p. 716 - 727.
- Matsumoto, D (2001) 'Culture and Emotion', in Matsumoto, D (ed), *The Handbook of Culture and Psychology*, p.171-194. New York, Oxford University Press.
- McCann, R.M. and & Giles, H. (2007) 'Age differentiated communication in organizations: Perspectives from Thailand and the USA'. *Communication Research Reports*, 24, p.1-12.
- McCrum, R. and Cran, W. and MacNeil, R. (1992) *The Story of English*. London, Faber & Faber.
- McDonough, J. and McDonough, S. (1997) *Research Methods for English Language Teachers*. Great Britain, Arnold, a member of the Hodder Headline Group.
- McNamara, T. (2000) *Language Testing*. Oxford, Oxford University Press.
- McNamara, T. and Roever, C. (2006) *Language Testing: The Social Dimension*. London, Blackwell.
- Mead, G.H. (1934) *Mind, self and society*. Chicago, University of Chicago Press.
- Melchers, G. and Shaw, P. (2003) *The English Language Series: World Englishes*. London, Hodder Arnold.
- Merriam, S. (1998) *Qualitative research and case study applications in education; Revised and expanded from case study research in education*. San Francisco, C.A, Jossey-Bass publications.
- Meyer, M. and Wodak, R. (2009) *Methods of Critical Discourse Analysis*. London, Sage Publications.
- Mullen, K. (1978) 'Direct evaluation of rater and scale in oral proficiency interviews', *Language Learning*, v.28, p.301-308.
- Myers, P. and Giles, H. and Reid, S. A. and Nabi, R. (2008). 'Law enforcement encounters: The effects of officer accommodativeness and crime severity on interpersonal attributions are mediated by intergroup sensitivity'. *Communication Studies*.
- Nakatsuhara, F. (2008) 'Inter-interviewer variation in oral interview tests', *ELT Journal*, vol. 62, no.3, p. 266-275.

- Nelson, G.L (1997) 'How Cultural Differences Affect Written and Oral Communication: The case of Peer Response groups'. *New Directions for Teaching and Learning*, 70, p.77-84. Retrieved 21.09.06 ERIC (EJ547589).
- Nishida, H. (1999) 'A cognitive approach to intercultural communication based on schema theory'. *International Journal of Intercultural Relations*, 23, 753-777.
- Norris, S. (2004) *Analyzing Multimodal Interaction: A Methodological Framework*. New York, Routledge.
- Norton, B. (2000) *Identity and Language Learning*. England, Pearson Education Limited.
- Nunan, D. (1992) *Research Methods in Language Learning*. Cambridge, UK, Cambridge University Press.
- O'Loughlin, K. (2002) 'The impact of gender in oral proficiency testing', *Language Testing*, vol. 19, no. 2, p.169-192.
- O'Loughlin, K. (2007) 'An investigation into the role of gender in the IELTS oral interview', in Taylor, L. and Falvey, P. (eds), *Studies in Language Testing 19*, Cambridge, Cambridge University Press.
- Orbe, M (1998) *Constructing co-cultural theory*. Thousand Oaks, CA, Sage.
- O'Sullivan, B. (2000) 'Exploring gender and oral proficiency interview performance.' *System* 28, 1-14.
- O'Sullivan, B. and Weir, C. (2002) 'Using observation checklists to validate speaking-test tasks', *Language Testing*, vol.19, no.1, p.33-56.
- O'Sullivan, B. and Yang, Lu. (2002) 'The impact of candidate language of examiner deviation from a set interlocutor frame in the IELTS Speaking Test.' *IELTS Research Reports*, Volume 6, 2006, p. 91-115. U.K, British Council, Australia, IELTS Australia Pty Limited.
- O'Sullivan, B. (2008) *Modelling Performance in Tests of Spoken Language*. Germany, Deutsche Nationalbibliothek.
- Paltridge, B. and Phakiti, A. (2010) *Continuum Companion to Research Methods in Applied Linguistics*. London, Continuum International Publishing Group.
- Parker, L. (1998) 'Race is...race ain't: An explanation of the utility of critical race theory in qualitative research in education.' *International Journal of Qualitative Studies in Education*, 11, 43-57.

- Patton, M. (2002) *Qualitative Research & Evaluation Methods*. London, Sage Publications Inc.
- Paul, I. (2004) 'IELTS as a predictor of academic language performance, part 2', *IELTS Research Reports*, Volume 7, 2007, p. 205-240. U.K, British Council, Australia, IELTS Australia Pty Limited.
- Pennycook, A. (1994). *The Cultural Politics of English as an International Language*. London and New York, Longman.
- Pennycook, A. (1998) *English and the Discourses of Colonialism*. Abingdon, Oxon, Routledge.
- Pennycook, A. (2007). *Global Englishes and Transcultural flows*. London and New York, Routledge.
- Philippot, P. and Feldman, R. and Coats, E. (ed) (1999) *The Social Context of Nonverbal Behavior*. Cambridge, UK, Cambridge University Press.
- Philips, S. (1972) 'Participant structures and communicative competence: Warm Springs children in community and classroom', in Cazden, C. and Hymes, D. (eds) *Functions of language in the classroom*, p.370-394. New York, Teachers College Press.
- Phillipson, R. (1992) *Linguistic Imperialism*. Oxford, Oxford University Press.
- Phillipson, R. (2001) 'English for globalisation or for the world's people?' *International Review of Education*, 47(3/4), p.185-200.
- Phillipson, R. (2008). 'The linguistic imperialism of neoliberal empire'. *Critical Inquiry in Language Studies*, 5(1), p.1-43.
- Piller, I. (2009) 'UAE Centre for Bilingualism and Bilingual Education (UAE CBBE)'. *International Journal of Bilingual Education and Bilingualism*, Vol 12 No. 2, March, 2009, p. 259 – 260.
- Quirk, R (1985) 'The English language in a global context' in Quirk R and Widdowson H.G. (ed) *English in the world: teaching and learning the language and literatures*. Cambridge, Cambridge University Press for the British Council.
- Quirk, R. (1990) 'What is Standard English?' *English in Use*, edited by Randolph Quirk and Gabriele Stein. London, Longman, p. 112-25.
- Read, J. and Nation, P. (2002) 'An investigation of the lexical dimension of the IELTS Speaking Test' in *IELTS Research Reports*, Volume 6, 2006, p. 207-231. U.K, British Council, Australia, IELTS Australia Pty Limited.

- Richards, B. and Malvern, D. (2000) 'Accommodation in oral interviews between foreign language learners and teachers who are not native speakers'. *Studia Linguistica*, 54, 260-271.
- Ridge, N (2010) 'Teacher Quality, Gender and Nationality in the United Arab Emirates: A crisis for boys'. *Dubai School of Government: working paper No. 10-06*, Dubai School of Government.
- Riley, P. (2007) *Language, Culture and Identity*. London, Continuum International Publishing Group
- Risager, K. (2007) *Language and Culture Pedagogy*. Clevedon, England, Multilingual Matters Ltd.
- Robertson, K. and Murachver, T. (2003) 'Children's speech accommodation to gendered language styles', *Journal of Language and Social Psychology*, vol. 22, no.3, p.321.
- Rock, P. (1979) *The making of symbolic interactionism*. Totowa, NJ, Rowman and Littlefield.
- Romano, J. and Bellack, A. (1980) 'Social validation of a component model of assertive behavior', *Journal of Consulting and Clinical Psychology*, vol.48, issue 4, August 1980, p.478-490.
- Rossner, R. and Bolitho, R. (ed) (1990) *Currents of Change in English Language Teaching*. Oxford, Oxford University Press.
- Ross, S. and Berwick, R. (1992) 'The discourse of accommodation in oral proficiency interviews', *Studies in Second Language Acquisition*, vol.14, p. 159 – 76.
- Rowley, J (2002) 'Using Case studies in Research'. *Management Research News*. Volume 25, Number 1 2002.
- Saayegh, F (2008) 'How can we maintain a national identity?', *The Gulf News*, May 27<sup>th</sup>, 2008.
- Salili, F. and Hoosain, R. (ed) (2007) *Culture, Motivation and Learning: A Multicultural Perspective*. Charlotte, NC, Information Age Publishing Inc.
- Sartor, V. and Heng, Y. (2008) 'World Englishes and Linguicism: the politics of language in China'. *ESL magazine*, November 2008, Keyways Publishing Ltd.
- Schaffner, C. (ed) (2000) *Translation in the Global Village*. Clevedon, England, Multilingual Matters Ltd.
- Schreier, D. et al (ed) (2010) *The Lesser-Known Varieties of English: An Introduction*. New York, Cambridge University Press.

- Schrober, M. and Conrad, F and Fricker, S. (2004) 'Misunderstanding standardized language in research interviews', *Applied Cognitive Psychology*, vol.18, no.2, p.169-188.
- Schuster, K. and Witkosky, D. (ed) (2007) *Language of the Land: Policy, Politics, Identity*. Charlotte, NC, Information Age Publishing Inc.
- Scollon, R. and Scollon, S. (1995) *Intercultural communication: a discourse approach*. Oxford, UK, Blackwell.
- Seedhouse, P and Egbert, M. (2004) 'The interactional organization of the IELTS speaking test', in *IELTS Research Reports*, Volume 6, 2006, p. 161-206. U.K, British Council, Australia, IELTS Australia Pty Limited.
- Sercu, L. (2004) 'Assessing intercultural competence: a framework for systematic test development in foreign language education and beyond'. *Intercultural Education*, 15 (1), p. 73-89, retrieved 12.09.06 ERIC (EJ694734).
- Sercu, L et al (2005) *Foreign Language Teachers and Intercultural Competence: An International Investigation*. Clevedon, England, Multilingual Matters Ltd.
- Shah, S (2004) 'The researcher in an intercultural context: a social intruder!' *British Educational Research Journal*, 30 (4), p. 549 – 575, retrieved 21.09.06. ERIC (EJ680377).
- Shalin, D.N. (1986) 'Pragmatism and social interactionism'. *American Sociological Review*, 51, p. 9-29.
- Shohamy, E. (1983) 'The stability of oral language proficiency assessment in the oral interview testing procedure', *Language Learning*, vol. 33, p.527 – 40.
- Shohamy, E. (1988) 'A proposed framework for testing the oral language of second/foreign language learners'. *Studies in Second Language Acquisition*, vol.10, p.165 – 79.
- Shohamy, E. (1994) 'The validity of direct versus semi-direct oral tests'. *Language Testing*, vol.11, p.99-123.
- Signorini, P. and Wiesemes, R. and Murphy, R. (2009) 'Developing alternative frameworks for exploring intercultural learning: a critique of Hofstede's cultural difference model', *Teaching in Higher Education*, Vol.14, No3, June 2009, 253-264.
- Sik, Hung Ng. (2007) 'From language acculturation to communication acculturation: addressee orientations and communication brokering in

conversations'. *Journal of Language and Social Psychology*, vol 26, no1, p.75 – 90.

Simpson, J. (2006) 'Differing expectations in the assessments of speaking skills of ESOL learners', *Linguistics and Education*, vol.17, no.1, p.40 – 55.

Slama-Cazacu, T (1976) Nonverbal components in message sequence: 'Mixed syntax'. In W.C McCormack & S.A. Wurm (Eds). *Language and man: Anthropological issues* (pp. 217 – 227). The Hague: Mouton.

Spolsky, B. (1989) *Conditions for Second Language Learning*. Oxford, Oxford University Press.

Stake, R. (1995) *The Art of Case Study Research*. Thousand Oaks, CA, Sage.

Stevick, E. (1982) *Teaching and Learning Languages*. Cambridge, Cambridge University Press.

Straw, J. (2006) cited in 'Deconstructing the veil', *Economist*, 381.8499. EBSCO. Web.

Stricker, L. (2004) 'The performance of native speakers of English and ESL speakers on the computer-based TOEFL and GRE General test', *Language Testing*, vol.21, no.2, p.146 – 173.

Swain, M. (1985) 'Large-scale communicative language testing: a case study', *New Directions in Language Testing*. Oxford, Pergamon Press.

Syed, Z, and Coombe, C. and Troudi, S. (ed) (2003) *Critical Reflection and Practice*. Dubai, UAE, TESOL Arabia.

Tannen, D (1990) *You just don't understand: Women and men in conversation*. New York, William Morrow.

Taylor, L. and Falvey, P. (ed) (2007) *IELTS Collected Papers (Research in Speaking & Writing Assessment)*. Cambridge, UK, Cambridge University Press.

Taylor, L. and Falvey, P. (ed) (2007) *IELTS Collected Papers: Studies in Language Testing*. Cambridge, UK, Cambridge University Press.

Taylor, S. and Wetherell, M. and Yates, S. (2009) *Discourse as Data*. London, Sage Publications.

Thakerar, J. and Giles, H. and Cheshire, J. (1982) 'Psychological and linguistic parameters of speech accommodation theory', in Fraser, C. and Scherer, K. (eds) *Advances in the Social Psychology of Language*, Cambridge, Cambridge University Press.

- Thompson, D (ed) (1995) 9<sup>th</sup> edition *Concise Oxford Dictionary*. Oxford: Oxford University Press.
- Thomson, R. (2006) 'The effect of topic of discussion on gendered language in computer-mediated communication discussion', *Journal of Language and Social Psychology*, vol.25, no.2, p.167-178.
- Thorpe, M. and Edwards, R. and Hanson, A. (ed) (1993) *Culture and Processes of Adult Learning*. London, Routledge.
- Ting-Toomey, S. (2005) 'Identity Negotiation Theory: Crossing Cultural Borders', in Gudykunst, W (ed) *Theorising about Intercultural Communication*, p.211-234, U.S.A, Sage Publications.
- Troudi, S. and Coombe, C. and Riley, S., (ed) (1999) *Unity Through Diversity*. Dubai, UAE, TESOL Arabia.
- Troudi, S. and Riley, S. and Coombe, C. (ed) (2002) *EFL Challenges in the New Millenium*. Dubai, UAE, TESOL Arabia.
- Trudgill, P. and Hannah, J. (1985) *International English* (second edition). London, Edward Arnold.
- Trudgill, P. and Hannah, J. (1994) *International English* (third edition). London, Edward Arnold.
- Tsui, A. (1996) 'Reticence and anxiety in second language learning', in Bailey, K. and Nunan, D. (eds) *Voices from the Language classroom: Qualitative Research in Second Language education*, Cambridge, Cambridge University Press.
- Tulloh, R. (1999) *IELTS Research reports Vol 2*. Canberra, National Library of Australia.
- Underhill, N. (2003) *Testing Spoken Language*. Cambridge, UK, Cambridge University Press.
- Valdes, J. (ed) (1986) *Culture Bound*. Cambridge, UK, Cambridge University Press.
- Van Leeuwen, T. (2008) *Introducing Social Semiotics*. Abingdon, Routledge.
- Vygotsky, L. (1986) *Thought and Language*. Cambridge, Massachussetts, The MIT Press.
- Walford, G. (ed) (1991) *Doing Educational Research*. London, Routledge.



Wallace, C. (1997) 'IELTS: global implications of curriculum and material design.' *ELT Journal* Volume 51/4, p. 370 – 373.

Wallace, C. (2002) 'Local literacies and global literacy' in Block, D and Cameron, D (eds) *Globalization and language teaching*, London, Routledge.

Warne, A. et al (ed) (2006) *Action Research in English Language Teaching in the UAE*. Abu Dhabi, UAE, HCT Press.

Watson, B. and Gallois, C. (1998) 'Nurturing communication by health professionals towards patients: A communication accommodation theory approach'. *Health Communication*. 10, p.343-355.

Wei, L. (ed) (2000) *The Bilingualism Reader*. London, Routledge.

Weir, C. (2005) *Language Testing and Validation*. Basingstoke, Palgrave MacMillan.

Weir, C. and Wu, J. (2006) 'Establishing test form and individual task comparability: a case study of a semi-direct speaking test', *Language Testing*, vol 23. no. 2, p.167-197.

Wengraf, T. (2001) *Qualitative Research Interviewing*. London, Sage Publications Ltd.

Wetherell, M. and Taylor, S. and Yates, S. (2001) *Discourse as Data: A Guide for Analysis*. London, England, Sage Publications Ltd. and The Open University Press.

Wetherell, M. and Taylor, S. and Yates, S. (2001) *Discourse Theory and Practice*. London, England, Sage Publications Ltd. and The Open University Press.

Wodak, R. and Meyer, M. (ed) (2009) *Methods of Critical Discourse Analysis (Second Edition)*. London, Sage Publication Ltd.

Woods, S. (1998) *IELTS Research reports Vol 1*. Canberra, National Library of Australia.

Xiaoming, X. (2007) 'Evaluating analytical scoring for the TOEFL Academic Speaking Test for operational use', *Language Testing*, vol.24, no.2. p. 251 – 286.

Yin, R (1994) *Case study research: design and methods*. 2<sup>nd</sup> edition. Thousand Oaks, CA: Sage.

Yule, G. (1985) *The Study of Language*. Cambridge, UK, Cambridge University Press.

Yoshida, R. (2001) 'Accommodation in Oral Proficiency Interviews'. Paper presented at the PAC3 conference November 22-25, 2001. Kitakyushu, Japan.

[www.ets.org/toefl](http://www.ets.org/toefl)

[www.ielts.org](http://www.ielts.org)



# APPENDICES

Appendix 1	IELTS public banding descriptors	iii
Appendix 2	HCT Band descriptors	iv
Appendix 3	Student permission form	vii
Appendix 4	Interview sample materials	viii
Appendix 5	Transcript of interview (sample)	xv
Appendix 6	Student questionnaire	xxv
Appendix 7	Faculty questionnaire	xxxiii
Appendix 8	Focus Group transcription (faculty excerpt)	xli
Appendix 9	Focus Group transcription (student excerpt)	xlii

## Appendix 1

### IELTS public banding scales



Band	Fluency and coherence	Lexical resource	Grammatical range and accuracy	Pronunciation
9	<ul style="list-style-type: none"> <li>speaks fluently with only rare repetition or self-correction; any hesitation is content-related rather than to find words or grammar</li> <li>speaks coherently with fully appropriate cohesive features</li> <li>develops topics fully and appropriately</li> </ul>	<ul style="list-style-type: none"> <li>uses vocabulary with full flexibility and precision in all topics</li> <li>uses idiomatic language naturally and accurately</li> </ul>	<ul style="list-style-type: none"> <li>uses a full range of structures naturally and appropriately</li> <li>produces consistently accurate structures apart from 'slips' characteristic of native speaker speech</li> </ul>	<ul style="list-style-type: none"> <li>uses a full range of pronunciation features with precision and subtlety</li> <li>sustains flexible use of features throughout</li> <li>is effortless to understand</li> </ul>
8	<ul style="list-style-type: none"> <li>speaks fluently with only occasional repetition or self-correction; hesitation is usually content-related and only rarely to search for language</li> <li>develops topics coherently and appropriately</li> </ul>	<ul style="list-style-type: none"> <li>uses a wide vocabulary resource readily and flexibly to convey precise meaning</li> <li>uses less common and idiomatic vocabulary skillfully, with occasional inaccuracies</li> <li>uses paraphrase effectively as required</li> </ul>	<ul style="list-style-type: none"> <li>uses a wide range of structures flexibly</li> <li>produces a majority of error-free sentences with only very occasional inappropriacies or basic/non-systematic errors</li> </ul>	<ul style="list-style-type: none"> <li>uses a wide range of pronunciation features</li> <li>sustains flexible use of features, with only occasional lapses</li> <li>is easy to understand throughout; L1 accent has minimal effect on intelligibility</li> </ul>
7	<ul style="list-style-type: none"> <li>speaks at length without noticeable effort or loss of coherence</li> <li>may demonstrate language-related hesitation at times, or some repetition and/or self-correction</li> <li>uses a range of connectives and discourse markers with some flexibility</li> </ul>	<ul style="list-style-type: none"> <li>uses vocabulary resource flexibly to discuss a variety of topics</li> <li>uses some less common and idiomatic vocabulary and shows some awareness of style and collocation, with some inappropriate choices</li> <li>uses paraphrase effectively</li> </ul>	<ul style="list-style-type: none"> <li>uses a range of complex structures with some flexibility</li> <li>frequently produces error-free sentences, though some grammatical mistakes persist</li> </ul>	<ul style="list-style-type: none"> <li>shows all the positive features of Band 6 and some, but not all, of the positive features of Band 8</li> </ul>
6	<ul style="list-style-type: none"> <li>is willing to speak at length, though may lose coherence at times due to occasional repetition, self-correction or hesitation</li> <li>uses a range of connectives and discourse markers but not always appropriately</li> </ul>	<ul style="list-style-type: none"> <li>has a wide enough vocabulary to discuss topics at length and make meaning clear in spite of inappropriacies</li> <li>generally paraphrases successfully</li> </ul>	<ul style="list-style-type: none"> <li>uses a mix of simple and complex structures, but with limited flexibility</li> <li>may make frequent mistakes with complex structures, though these rarely cause comprehension problems</li> </ul>	<ul style="list-style-type: none"> <li>uses a range of pronunciation features with mixed control</li> <li>shows some effective use of features but this is not sustained</li> <li>can generally be understood throughout, though mispronunciation of individual words or sounds reduces clarity at times</li> </ul>
5	<ul style="list-style-type: none"> <li>usually maintains flow of speech but uses repetition, self-correction and/or slow speech to keep going</li> <li>may over-use certain connectives and discourse markers</li> <li>produces simple speech fluently, but more complex communication causes fluency problems</li> </ul>	<ul style="list-style-type: none"> <li>manages to talk about familiar and unfamiliar topics but uses vocabulary with limited flexibility</li> <li>attempts to use paraphrase but with mixed success</li> </ul>	<ul style="list-style-type: none"> <li>produces basic sentence forms with reasonable accuracy</li> <li>uses a limited range of more complex structures, but these usually contain errors and may cause some comprehension problems</li> </ul>	<ul style="list-style-type: none"> <li>shows all the positive features of Band 4 and some, but not all, of the positive features of Band 6</li> </ul>
4	<ul style="list-style-type: none"> <li>cannot respond without noticeable pauses and may speak slowly, with frequent repetition and self-correction</li> <li>links basic sentences but with repetitious use of simple connectives and some breakdowns in coherence</li> </ul>	<ul style="list-style-type: none"> <li>is able to talk about familiar topics but can only convey basic meaning on unfamiliar topics and makes frequent errors in word choice</li> <li>rarely attempts paraphrase</li> </ul>	<ul style="list-style-type: none"> <li>produces basic sentence forms and some correct simple sentences but subordinate structures are rare</li> <li>errors are frequent and may lead to misunderstanding</li> </ul>	<ul style="list-style-type: none"> <li>uses a limited range of pronunciation features</li> <li>attempts to control features but lapses are frequent</li> <li>mispronunciations are frequent and cause some difficulty for the listener</li> </ul>
3	<ul style="list-style-type: none"> <li>speaks with long pauses</li> <li>has limited ability to link simple sentences</li> <li>gives only simple responses and is frequently unable to convey basic message</li> </ul>	<ul style="list-style-type: none"> <li>uses simple vocabulary to convey personal information</li> <li>has insufficient vocabulary for less familiar topics</li> </ul>	<ul style="list-style-type: none"> <li>attempts basic sentence forms but with limited success, or relies on apparently memorised utterances</li> <li>makes numerous errors except in memorised expressions</li> </ul>	<ul style="list-style-type: none"> <li>shows some of the features of Band 2 and some, but not all, of the positive features of Band 4</li> </ul>
2	<ul style="list-style-type: none"> <li>pauses lengthily before most words</li> <li>little communication possible</li> </ul>	<ul style="list-style-type: none"> <li>only produces isolated words or memorised utterances</li> </ul>	<ul style="list-style-type: none"> <li>cannot produce basic sentence forms</li> </ul>	<ul style="list-style-type: none"> <li>speech is often unintelligible</li> </ul>
1	<ul style="list-style-type: none"> <li>no communication possible</li> <li>no rateable language</li> </ul>			
0	<ul style="list-style-type: none"> <li>does not attend</li> </ul>			

## Appendix 2

### HCT Speaking band descriptors

Band	Communicative range x 3	Overall fluency x 1	Accuracy and appropriacy x 1	Pronunciation, intonation and stress x 1	Raw score
9	<p>Can converse relevantly and interestingly on a wide range of academic and non academic topics both inside and outside own sphere with other educated native speakers.</p> <p>Uses a wide range of cohesive and sequencing devices in extended discourse accurately and appropriately</p>	Fluency comparable to that of an educated native speaker, with total flexibility to adapt to change of topic and conversational context	Wide range of vocabulary and idiom, stylistic and structural language features are used appropriately and accurately	Only a slight accent may be noticeable which in no way intrudes on the communication. Intonation and stress patterns comparable with native speaker.	54
8	<p>Can converse effectively on all academic and non academic topics which relate to own experience and interests. Can sustain well organised, extended discourse involving speculation, argumentation, description and narration. Meaning is precisely conveyed</p>	Converses fluently with no barrier to communication.. Flexible enough to adapt to change of topic and conversational context	Accurate use of a wide range of linguistic features, including complex sentences, cohesive devices, and modifiers. Only occasional slight inappropriacies or 'slips of the tongue' may occur	Accent may be noticeable, but does not affect communication. Intonation and stress patterns approach native speaker level.	48
7	<p>Can communicate competently on a wide range of academic and non academic topics which relate to own experience and interests. Extended discourse is organised and suitable cohesion used. Can use abstract speculative and argumentative language effectively in some, but not all, contexts. Meaning is clearly conveyed.</p>	Generally fluent, with occasional pauses for repair, circumlocution or 'searching'. Can cope with switches of topic and adapt to some extent to changes in conversational context.	Fairly accurate use of a wide range of linguistic features, including complex sentences, cohesive devices, and modifiers. Some linguistic errors, and lexical inappropriacies, but meaning is not impaired.	Intonation appropriate. Fully intelligible. L1 stress patterns and accent may be noticeable but no strain is felt in communication.	42

Band	Communicative range x 3	Overall fluency x 1	Accuracy and appropriacy x 1	Pronunciation, intonation and stress x 1	Raw score
6	Can communicate satisfactorily on general, vocational and social topics. Can present information confidently but abstract speculation and argument may break down under pressure. Can deal with longer and more complex description and narration.	Hesitations rare, usually only when searching for unfamiliar vocabulary. Flexible enough to cope with topic switches.	Can form complex sentences reasonably accurately. Vocabulary adequate to express finer meaning and to modify.	Intonation appropriate. Fully intelligible. L1 stress patterns and accent are noticeable but no strain is felt in communication.	36
5	Can initiate and sustain extended conversation on general topics. Cannot handle abstract topics, argument or speculation. Recognises and expresses attitude, though not finely. Can elicit and leave a telephone message satisfactorily	Hesitations occur as speaker searches for vocabulary, but repair strategies are developed enough to keep the conversation going at near normal speed.	Basic sentence structure usually correct, but errors occur in more complex sentences. Basic tenses used appropriately. Vocabulary inadequate for abstract discussion or fine description	Mainly intelligible. Uses intonation appropriate to the context. Interlocutor may occasionally request repetition and clarification.	30
4	Can initiate and sustain conversation on everyday topics.  Can give and elicit simple information face to face and, in a more limited way, on the telephone. Can describe a simple process or series of events	Shows more spontaneity when on familiar ground, but hesitations still occur and rephrasing and prompting is still needed.	Can use basic sentence and question forms, although with many inaccuracies. Cannot form complex sentences. Can use basic sequencers. Vocabulary sufficient for everyday and work related needs.	Reasonably comprehensible to native speakers, though  interlocutor may feel some strain, and misunderstanding may still occur.	24
3	Can engage in basic communication on familiar personal or work related topics. Can give simple description or instructions. Can elicit only the basic (name, number) elements of a telephone message.	Responds to rather than initiates conversation, but can ask for repetition and express lack of understanding. Still requires a patient interlocutor	Evidence of basic sentence structure and question forms starting to appear, although inaccurate. Vocabulary can only convey basic information	Misunderstanding still occurs, due to L1 interference, stress and intonation despite effort from the listener	18
2	Can respond to simple instructions and questions in face to face situations	Frequent hesitations, but utterance length increases. Requires support and tolerance from interlocutor in the form of repetition and rephrasing	Basically unable to form sentences. Can question only by using intonation. Formulaic expressions may be accurate.	Can be understood with some difficulty by instructors and is occasionally unintelligible	12
1	Can respond to the simplest requests for personal information. Can use one or two basic courtesy formulae such as greetings. Can ask for things by pointing or gesture	Limited to a few words, or short utterances with frequent hesitation. Totally dependent on interlocutor to repeat, paraphrase and prompt	Odd words and formulaic phrases may be used appropriately	Can be understood only with difficulty, even by instructors. Frequently unintelligible	6

## GUIDE FOR ASSESSORS

The descriptors are weighted heavily towards the communicative range descriptor. This is intended to give a global picture of what the candidate can do. The nature of the assessment may automatically limit the assessor to within one or two communicative range bands. e.g. If the assessment instrument requires the candidate only to ask and give personal information, then the communicative range expected will be 4. The other elements may modify this assessment up or down.

A profile of an above average candidate may be as follows:

Task : Describe how use a photocopier

- Communicative range = 4 (weighted x 3) = 12
- (i.e. candidate described the process satisfactorily using linkers)  
Overall fluency = 5 (weighted x 1) = 5
- Accuracy and appr. = 5 (weighted x 1) = 5
- Pron. int. & stress = 5 (weighted x 1) = 5

Total 27

Divide by 6 (combined weighting) Final band **4.5**



### Appendix 3

#### **Guardian's Permission to Photograph and/or Video**

**Student Name:**

**Student ID:**

**Section:**

Dear Parent / Guardian

We would like to inform you that Sharjah Women's College students will be interviewed by an English teacher, Mrs. Gail AlHafidh.

The interview will be a practice IELTS interview and will form part of a doctoral research study into intercultural communication. Any video material generated will remain strictly confidential.

**Guardians:**

I / \_\_\_\_\_

The Guardian of / \_\_\_\_\_

ID No. / \_\_\_\_\_

Section / \_\_\_\_\_

Guardian Tel. No./ \_\_\_\_\_

**have no objection to the above request.**

Guardian's  
Relationship: (Father/Husband/Guardian)

Guardian's Signature: \_\_\_\_\_

Date: \_\_\_\_\_

***Note: This form must be signed by the student's guardian only. Violation of this might lead to the student's dismissal from the College.***

**Evan Jones  
Chair of English**

#### **تصريح ولي الأمر بإجراء تصوير فوتوغرافي و / أو سينمائي**

اسم الطالبة :

الرقم الجامعي:

الفصل:

السيد / ولي الأمر الخريجة المعترم

نفيدكم علماً بأننا سنقوم بإجراء مقابلات مصورة لامتحان اللغة الانجليزية مع طالبات كليات الشارقة التقنية العليا للبنات من قبل معلمتهن السيدة/ جايل الحفيد. هذا و ستكون المادة التصويرية للمقابلة جزءاً من بحث علمي في مجال الاتصال الثقافي. و نضمن لكم خصوصية جميع المواد المصورة. الرجاء التوقيع أدناه في حال موافقتكم.

**ولي الأمر:**

أنا / \_\_\_\_\_

ولي أمر الطالبة / \_\_\_\_\_

رقمها الجامعي / \_\_\_\_\_

في قسم / \_\_\_\_\_

هاتف ولي الأمر : \_\_\_\_\_

لا مانع لدي حسب الطلب المذكور أعلاه.

صلة القرابة : ( الأب / الزوج / ولي الأمر):

توقيع ولي الأمر: \_\_\_\_\_

التاريخ : \_\_\_\_\_

***ملحوظة هامة: على الطالبة توقيع هذا الإذن من قبل ولي الأمر فقط وأي مخالفة لهذا الأمر قد تعرض الطالبة لعقوبة قد تصل إلى حد الفصل من الكلية.***

**فيل كوركوس  
عميد قسم اللغة الانجليزية و التربية**

## Sample Interview Questions

There are 3 parts to the interview:

- General Qs (4-5 mins)
- Topic (3-4 mins)
- Extended Qs based on topic (4-5 mins)

### Part 1

#### **Introductory Frame (30 secs)**

Good morning/afternoon

I'm .....

Can you tell me your full name, please?

Thank you, and what shall I call you?

Can I see your identification, please?

Thank you, that's fine.

Now, in this first part I'd like to ask you a few questions about yourself.

#### **Frame 1 (4-4.5 mins) Choose from Frame 1 or 2 and then 2 more of your choice**

Let's talk about your hometown or village

- Where exactly do you come from?
- Can you describe it to me?
- How many people live there?
- What kinds of jobs do people do there?
- Do you like the area? Why?

OR

#### **Frame 2**

Let's talk about your studies

- What program are you in?
- Why did you choose this course?
- Tell me about the different subjects you study
- What do you like best about the course?
- What job opportunities are there in this field?

#### **Frame 3**

Let's talk about your friends

- Where did you meet most of your friends?
- How do you spend time with them?
- Do you see them outside college?
- Do you think that you will still be friends in 20 years time?

#### Frame 4

Let's talk about holidays

- How do you usually spend the holidays?
- Have you traveled abroad?
- Who do you usually go on holiday with?
- What would be your ideal holiday?

#### Frame 5

Let's talk about special events

- What are the important events in your culture?
- How do you celebrate them?
- Do you offer gifts? If so, what kind?
- Have you always celebrated these events?

#### Frame 6

Let's talk about music

- Do you listen to music often?
- What kind of music do you prefer?
- How important is music in your culture?
- Why do people listen to music?

#### Frame 7

Let's talk about sports

- Do you enjoy playing sports?
- Do you like watching sports?
- Why do people enjoy watching/playing sports?
- Do men and women play/watch the same kinds of sports in your country?

#### Frame 8

Let's talk about transport

- How do you usually travel to college?
- Is transport a problem in your country?
- Would you like to be a pilot or a train driver? Why?
- What would be your ideal way of traveling?

### Frame 9

Let's talk about seasons

- What are the main seasons in your country?
- Which is your favourite. Why?
- What is your favourite kind of weather? Why?
- What problems does the weather cause in your country?

### Frame 10

Let's talk about buildings

- Which building is the most important in your country, in your opinion?
- Do you like modern or traditional buildings?
- Is accommodation an issue in your country?
- Where would your dream house be located? Why?

### Frame 11

Let's talk about clothes and fashion

- Do you think it is important to wear fashionable clothes?
- Do you think children should wear fashionable clothes?
- Which is your favourite item of clothing? Why?
- Is fashion important to all ages, in your opinion?

### Frame 12

Let's talk about animals

- Do you have a pet?
- What animal is most popular in your country?
- Why is it so special?
- What animal do you dislike most? Why?

### Frame 13

Let's talk about food

- What is your favourite food?
- What are typical dishes in your country?
- Do you enjoy foreign food? Which?
- Who make the best cooks, men or women? Why?

## Frame 14

Let's talk about the media

- Do you listen to the radio? When?
- What is your favourite radio/TV programme?
- How do most people learn about the news?
- How important are newspapers?

## Interviewer's Card

### Part 2

#### Individual Long Turn (3-4 minutes)

Now I'm going to give you a topic and I'd like you to talk about it for a minute or two. Before you start, you'll have a minute to think about what you are going to say. You can make some notes if you wish. O.K?

Here's some paper and a pencil for making notes.  
....and here's your topic

I'd like you to describe your favourite advertisement.

*(Allow up to a minute for preparation, but the candidate can start earlier if he/she wants to)*

*N.B do not read this out!*

#### **Describe your favourite advertisement**

**You should say:**

**What product is being advertised**

**What media is used for the advertisement (TV/magazine etc)?**

**Explain why you like it so much.**

**Does the advertisement persuade you to buy the product?**

O.K? Remember you have 1 to 2 minutes for this so don't worry if I stop you. I'll tell you when the time is up. Can you start speaking now please?

*(At the end of the candidate's long turn ask one or both of the rounding off questions to close this part of the test)*

Does anyone else share your opinion?

Have you ever bought something just because you liked the advertisement?

## **PART 2**

**Describe your favourite advertisement**

**You should say:**

**What product is being advertised**

**What media is used for the advertisement (TV/magazine etc)?**

**Explain why you like it so much.**

**Does the advertisement persuade you to buy the product?**

You will have to talk about the topic for 1 to 2 minutes.

You have one minute to think about what you are going to say.

You can make some notes to help you if you wish.

### PART 3

Discussion (4-5 minutes)

Examiner: We've been talking about your favourite advertisement and I'd like to discuss with you one or two more general questions related to this. Let's consider first of all.....

*(Examiner to select one or two of the following bullet-pointed themes and use the sub-themes to develop the discussion)*

- **Advertisements in general**

Describe how advertisers persuade people to buy their products.  
Consider why people consumers would prefer one brand over another (e.g soap powder/shampoo)  
Speculate as to what advertisements will be like in the future.

- **Censorship/Media**

Consider the differences between advertisements in your culture and in other countries.  
Are there some products that should not be advertised in your opinion? Why?  
Should advertisements be censored by governments? Why?  
Should famous sportsmen and film stars feature in advertisements? If so, what is the effect of this on young people? (Positive or negative?)



# Appendix 5

## Sample Transcription of video interview

Int 4: M Assessor/fully veiled student 2.58 mins	Eye Movement	Head Movement	Facial Gestures	Posture	Proxemics	Hand Movements	Fidgeting	Other	Questions asked	Comments
	Looking down at file, looks up for 1 second only			leaning forward, hands below table	chair pushed away from the table, directly facing c		restless- ness, shuffling in chair slightly		<i>General start- up Qs Where exactly do you come from?</i>	
31-60	Looks up for 2 secs on 2 separate occasions...does not look up to ask the Qs		smile			hands held below the table		Sits back, up and then sinks back to leaning forward almost slumped over the file, shoulders hunched.	<i>Can you describe Ajman for me. How many people live in Ajman?</i>	
61-90	Very little eye contact, 1 sec at a time...not sustained, only briefly AFTER asking a Q						Shuffles chair back and forth. Fiddling with tie.		<i>What kind of jobs do people do there? How do you usually spend your holiday?</i>	

91-120		smile		slight tapping of knee, swing of chair side to side	Who do you usually go on holiday with? What would be your ideal holiday?
121-150	On repeating the Q, s sustains gaze for 2-3 secs	smile			Where would you go? How do you usually travel to college?
151-180	Only Q a looks at s to ask. Slightly longer periods of gaze. 2-3 secs	smile		shuffles chair back and forth.	Is transport a problem in this country? Would you like to be a pilot or a train driver? A clearly thinks this is a daft Q - why did he ask it?
181-210					
211-240					
241-270	General: very short turn, noticeable lack of attempts to look at C, tangible ill of ease				
271-330					
331-330					

Int 4: M Assessor/fully veiled student. 2 mins 58 secs		Eye Movement	Head Movement	Facial Gestures	Posture	Proxemics	Hand Movements	Fidgeting	Other	Questions asked	Comments
00-30	No eye contact visible - fully veiled				Sitting upright, shoulder s relaxed	arm's length away from table	Hands held together loosely on lap	Swinging the chair side to side slightly		<i>General start- up Qs Where exactly do you come from?</i>	
31-60							Frequent hand movement			<i>Can you describe Ajman for me. How many people live in Ajman?</i>	
61-90							Frequent hand movement, especially RH	Swinging the chair side to side slightly		<i>What kind of jobs do people do there? How do you usually spend your holiday?</i>	
91-120					Lean's forward to ask for Q to be repeated		Frequent hand movement, especially RH	Swinging the chair side to side slightly		<i>Who do you usually go on holiday with? What would be your ideal holiday?</i>	Has to repeat Q

121-150
151-180
181-210
211-240
241-270
271-330
331-330

Frequent  
hand  
movement,  
especially RH

Swinging  
the chair  
side to  
side  
slightly.  
Fiddles  
with veil

*Where would  
you go?How  
do you  
usually travel  
to college?*

Frequent  
hand  
movement,  
especially RH

Swinging  
the chair  
side to  
side  
slightly

*Is transport a  
problem in  
this  
country?Wou  
ld you like to  
be a pilot or a  
train driver?*

Int 4: M Assessor/fully veiled student. 2 mins 51 secs	Eye Moveme nt	Head Movement	Facial Gestures	Posture	Proxemics	Hand Movements	Fidgeting	Other	Questions asked	Comments
00-30									Now I'm going to give you a topic: I'd like you to describe your favourite image, picture of photograph.	
31-60							Swinging the chair side to side slightly	Lifts veil to look at the prompt card		
61-90 91-120						Frequent hand movement, especially RH	Swinging the chair side to side slightly			
121-150 151-180		Shakes head - no					Swinging the chair side to side slightly			
181-210 211-240 241-270 271-330 331-330										

**Int 4: M**  
**Assessor/fully**  
**veiled**  
**student. 2**  
**mins 51 secs**

Eye Movement	Head Movement	Facial Gestures	Posture	Proxemics	Hand Movements
-----------------	------------------	--------------------	---------	-----------	-------------------

Fidgeting

Other

**Questions  
asked**

**Comments**

00-30

Limited -  
twice for a  
second

Leaning  
forward,  
arms on  
legs. Sits  
backs and  
away as  
student lifts  
veil to read  
the prompt  
card.

*Now I'm  
going to  
give you a  
topic: I'd  
like you to  
describe  
your  
favourite  
image,  
picture of  
photograph.*

*Clearly  
uncomfortable,  
especially  
when veil is  
lifted....perhaps  
not sure of  
what is  
expected of  
him.*

31-60

Fiddles  
wit file  
while  
students  
is  
making  
notes

61-90

91-120

121-150

151-180

181-210

211-240

241-270

271-330

331-330

Int 4: M Assessor/fully veiled student. 4 mins 26 secs	Eye Movement	Head Movement	Facial Gestures	Posture	Proxemics	Hand Movements	Fidgeting	Other	Questions asked	Comments
00-30	Looking down at file, looks up for 1 second only			Leaning forward, hands below table	chair pushed away from the table, directly facing c		restless- ness, shuffling in chair slightly	Scratching leg	<i>Can you describe to me the types of events that people take photos of in your culture? Why do people take photographs of important events?</i>	
31-60	Looks up once									
61-90	longest period of eye contact				Uses hands to portray opening a photo album		Very fidgety, foot tapping, swaying back and forward slightly		<i>In the future do you think the way in which we record important events will change? Can you tell me how it will change? Where do you store your photographs?</i>	s/S needed support/repetition of question here - longest eye contact as support mechanism?

91-120		Moves head to denote past				Where do you store your photographs?
121-150	One sec look up to c after pause in conversation		Leaning forward, hands below table		Raises eyebrows quizically, Laughs. Scratches leg.	Can you describe any traditional art forms in your culture
151-180	Prolonged eye contact	Smile		Uses hands to act out writing calligraphy		has to prompt..what about calligraphy? Uses gesture and eye contact when student appears to be in distress/unable to answer
181-210	Quick look up after asking the Q 2 further glances up					How important is art here in school compared with science?
211-240	Looks at c to tell her "nearly finished". Sustained glance while he asks the q		Leans backwards and away from table, sits up	Raises hands to gesture graphic art	Tapping feet	To what extent do you think computer graphics can be considered art?
241-270						Has to mention we've nearly finished. Clearly aware of her unease.
271-330						
331-330						



Int 4: M Assessor/fully veiled student. 4 mins 26 secs		Eye Movement	Head Movement	Facial Gestures	Posture	Proxemics	Hand Movements	Fidgeting	Other	Questions asked	Comments
00-30		Looking directly at a			Sitting upright, facing a	arm's length away from table facing a	Frequent hand movement, especially RH	Swingin side to side on chair		<i>Can you describe to me the types of events that people take photos of in your culture? Why do people take photographs of important events?</i>	
31-60										<i>In the future do you think the way in which we record important events will change?</i>	s/S needed support/re petition of question here
61-90						Leans forwards slightly	Hand on table briefly as she asks for repetition of q				
91-120			Appears to look away/down				Hand on table briefly as she asks for repetition of q	Swinging side to side on chair		<i>Can you tell me how it will change?Where do you store your photographs?</i>	Appears to put hand on table when unsure of Q - trying to accommod ate to gain better understand ing/commu

121-150

151-180

181-210

211-240

241-270

271-330

331-330

Turns away  
slightly

Hands off the  
table when she  
feels O.K to  
speak. Hands  
gesticulating to  
'air' paint  
calligraphy.

Raises hands ,  
fingers  
outstretched  
signifying stress

Frequent hand  
movement,  
especially RH

both arms on  
desk

Laughs

touches  
face veil

*Can you  
describe any  
traditional art  
forms in your  
culture?*

*How  
important is  
art here in  
school  
compared with  
other  
subjects?*

*To what extent  
do you think  
computer  
graphics can  
be considered  
art?*

nication?

has to  
prompt..w  
hat about  
calligraphy  
?

*Has to  
mention  
we've  
nearly  
finished*

## Appendix 6

### Student Questionnaire

*This questionnaire forms part of my research into the oral (speaking) part of your English assessment at college. Please answer according to **your own experience** – there are no right or wrong answers, I am only interested in finding out about your thoughts and reactions to different parts of the speaking test. All of your responses will be completely anonymous (unnamed) and treated confidentially (not discussed with anyone).*

#### SECTION A (Background Information)

General Information Please circle your answer.

- 1) How old are you?  
i) 16-18 ii) 19-21 iii) 22+
- 2) What type of High School did you attend?  
i) Government ii) Private iii) Both
- 3) What was the main language used for teaching?  
i) English ii) Arabic iii) Other
- 4) What gender were your teachers?  
i) Male only ii) Female only iii) Both male & female teachers
- 5) Was English the first language of your last High School English teacher?  
i) Yes ii) No.....if no, please state the language \_\_\_\_\_
- 6) What is your current program major at HCT?  
i) Business ii) I.T iii) Engineering iv) Education v) Health vi) Applied Media

#### Language learning context

- 7) Is Arabic your '*mother tongue*' (main language spoken at home)?  
i) Yes ii) No.....if no, please state the language \_\_\_\_\_
- 8) What OTHER languages can you speak and to what level? (e.g F (fluent) A (advanced) G (good) B (basic) )  
\_\_\_\_\_  
\_\_\_\_\_
- 9) What countries, outside the U.A.E, have you visited for a month or longer.  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_
- 10) Where does your HCT English teacher come from? Please circle your answer.  
a. Britain/Ireland b. America/Canada c. Australia/N. Zealand d. S. Africa e. Don't know

11) How important is it to you that your English teacher is a "native" speaker? (English is their first language).

*Please circle your answer.*

a. Very important b. Important c. Unimportant d. Not sure

Explain your answer

---

---

12) Do you feel that your English teacher has a basic understanding of Arabic language, and traditions? *Please circle your answer.*

a. Yes b. No c. Not sure

Explain your answer

---

---

13) How important is it to you that your English teacher has a basic understanding of Arabic language and traditions? *Please circle your answer.*

a. Very important b. Important c. Unimportant d. Not sure

Explain your answer

---

---

**SECTION B** (College speaking test -IELTS preparation).

In **part 1** of the college SPEAKING test, the topics are usually about your family, home background and free time activities.

- 14) How willing are you to talk about **these topics** with the examiner, a person who is outside your circle of family & friends? Please complete the table below by marking an **X** in the appropriate box. Add comments to explain your answer if you can.

Topic	Very willing	O.K	Not very willing	Not at all willing	Don't know	Comments
<b>Your family</b> (for example: number and ages of sisters/brothers, types of jobs they do, who lives with you.....)						
<b>Your friends</b> (for example: do you prefer lots of friends or just a few, what are the important qualities of a friend, describe your best friend.....)						
<b>Yourself</b> (for example: your personality, favourite song, item of clothing, shops etc.....)						
<b>Your home</b> (for example: is it a villa or an apartment, location, local facilities.....)						
<b>College life</b> (for example: daily routine, college rules, teachers, subjects.....)						
<b>Free time activities</b> (for example: what do you like to do when you are not at college, how often do you go to the cinema .....)						

In **part 2** of the college SPEAKING test you are asked to talk about a topic for between 1-2 minutes.

- 15) How willing are you to talk about **these topics** with the examiner, a person who is outside your circle of family & friends? Please complete the table below by marking an **X** in the appropriate box. Add comments to explain your answer if you can.

Topic	Very willing	O.K	Not very willing	Not at all willing	Don't know	Comments
Music						
Art						
T.V programs/films						
Photographs						
Books						
Flowers & plants						
International news & events						
Local news & events						
Celebrities						
Fashion						
A memorable day						
Religious festivals						
Travel abroad						
Weddings						
Transport						
A Hotel stay						
Concerts & plays						
Gift giving and receiving						
Famous Buildings						
Environment						

Assessment context

16) Do you feel that the interview room is an appropriate environment for the speaking assessment. *Please circle your answer*

- a. Yes      b. No      c. Not sure

Explain your answer

---

---

17) Do you have a preference for the gender of examiner? *Please circle your answer*

- a. Female      b. Male      c. No particular preference

Explain your answer

---

---

18) What test arrangements would you prefer? *Please circle your answer*

- a. 1 student - 1 examiner  
b. 2 students - 1 examiner  
c. Group of students (3 or more): 1 examiner  
d. No particular preference

Explain your answer

---

---

19) Would you prefer to attend the speaking exam wearing a veil (partially or fully)?

- a. Yes      b. No      c. Not sure

Explain your answer

---

---

20) Is your decision to not wear a veil **or** to wear one (partially or fully) affected by the gender of the examiner?

Explain your answer

---

---

Communication during the speaking assessment

21) Do you think that covering your face, either fully or partially, would affect the quality of communication in the speaking assessment? *Please circle your answer*

- a. Yes, it would be better.
- b. Yes, it would be worse.
- c. No effects
- d. Not sure

Explain your answer

---

---

22) Do you think that **the examiner** would react positively or negatively to you covering your face, either fully or partially, in the speaking assessment? *Please circle your answer*

- a. Positively
- b. Negatively
- c. Indifferently (no reaction)
- d. Not sure

Explain your answer

---

---

23) Would **you** react positively or negatively to the examiner covering their face, either fully or partially, in the speaking assessment? *Please circle your answer*

- a. Positively
- b. Negatively
- c. Indifferently (no reaction)
- d. Not sure

Explain your answer

---

---



24) Would you prefer to be interviewed online/via telephone/via videoconferencing than in a face-to-face interview? *Please circle your answer*

a. Yes      b. No      c. Not sure

Explain your answer – is gender a factor or are there other reasons affecting your response?

---

---

**SECTION C** (Varieties of English)

25) Can you distinguish (notice the differences) between varieties of spoken English, such as British, American and Australian, for example? *Please circle your answer*

- a. Yes      b. No      c. Not sure

Explain your answer

---

---

26) Is it important to you that you speak, **during the exam**, in a particular way and that this may affect the grade outcome ? *Please circle your answer*

- a. Yes, British    b. Yes, American    c. Yes, Australian    d. No    e. Not sure

Explain your answer

---

---

27) Do you feel that it is important to the examiner that you speak, **during the exam**, in a particular way and that this may affect the grade outcome? *Please circle your answer*

- a. Yes, British    b. Yes, American    c. Yes, Australian    d. No    e. Not sure

Explain your answer

---

---

**THANK YOU SO MUCH FOR ANSWERING THIS QUESTIONNAIRE. ALL INFORMATION GIVEN WILL REMAIN CONFIDENTIAL(not shared with anyone) AND USED ONLY FOR THE PURPOSES OF THIS RESEARCH PROJECT..**

**Examiner/Faculty Questionnaire**

**EFL Assessment - speaking**

*This questionnaire forms one part of my research into EFL oral assessment. I would be very grateful if you would answer the following questions according to **your own experience** as an examiner/teacher of EFL. All responses will be completely anonymous and confidentiality is guaranteed.*

**SECTION A** (Background Information)

**General Information** *Please circle your answer.*

1) Gender

i) Female ii) Male

2) Age category:

i) 18-30 ii) 31-50 iii) 50+

3) Gender of students you currently teach

i) M ii) F iii) Both

4) Please state your country of origin (e.g Britain, Ireland)

5) Please state your 'Mother tongue' (e.g Arabic, English, Farsi, etc...)

6) Please state which other languages you speak and for each language, please indicate level e.g F (fluent) A (advanced) I (intermediate) B (beginner)

7) Please list the countries, outside the U.A.E, where you have worked and state for how long (approximately, in terms of months/years).

**Language teaching context: UAE**

8) How important, in the context of teaching English, do you feel is it to have a basic understanding of Arabic & local cultural traditions? *Please circle your answer.*

a. Very important b. Important c. Unimportant d. Not sure

Explain your answer

9) How important to your students do you think is it that you have a basic understanding of Arabic & local cultural traditions? *Please circle your answer.*

a. Very important   b. Important   c. Unimportant   d. Not sure

Explain your answer

---

---

10) How important to your students do you think is it that you are a native English speaker? *Please circle your answer.*

b. Very important   b. Important   c. Unimportant   d. Not sure

Explain your answer

---

---

## **SECTION B** (College speaking test -IELTS preparation).

In **part 1** of the college SPEAKING test, the topics are usually general and are about family, home background and free time activities.

- 11) How willing are you to talk about **these topics** with the student? Please complete the table below by marking an **X** in the appropriate box. It would be very helpful if you could add comments to explain your answer.

Topic (asked of student)	Very willing	O.K	Not very willing	Not at all willing	Don't know	Comments
<b>Your family</b> (for example: number and ages of sisters/brothers, types of jobs they do, who lives with you.....)						
<b>Your friends</b> (for example: do you prefer lots of friends or just a few, what are the important qualities of a friend, describe your best friend.....)						
<b>Yourself</b> (for example: your personality, favourite song, item of clothing, shops etc.....)						
<b>Your home</b> (for example: is it a villa or an apartment, location, local facilities.....)						
<b>College life</b> (for example: daily routine, college rules, teachers, subjects.....)						
<b>Free time activities</b> (for example: what do you like to do when you are not at college, how often do you go to the cinema .....)						

In **part 2** of the college SPEAKING test the student is asked to talk about a topic for between 1-2 minutes.

- 12) In your experience, how comfortable would the students be to talk about the topics listed below. Please complete the table by marking an X in the appropriate box. It would be very helpful if you could add comments to explain your answer.

Topic	Very comfortable	O.K	Not very comfortable	Zero comfort	Don't know	Comments
Music						
Art						
T.V programs/films						
Photographs						
Books						
Flowers & plants						
International news & events						
Local news & events						
Celebrities						
Fashion						
A memorable day						
Religious festivals						
Travel abroad						
Weddings						
Transport						
A Hotel stay						
Concerts & plays						
Gift giving and receiving						
Famous Buildings						
Environment						

Assessment context:

13) Do you feel that the interview room, in this context, is an appropriate environment for the speaking assessment? *Please circle your answer*

- a. Yes      b. No      c. Not sure

Explain your answer

---

---

14) Do you think that female students, in the UAE cultural context, would have a preference as to the gender of the examiner for speaking exams, given a choice? *Please circle your answer*

- a. Yes, female    b. Yes, male    c. No preference    d. Not sure

Explain your answer

---

---

15) Do you think the students would have a preference to the student:examiner ratios? If so, which scenario do you think they might prefer? *Please circle your answer*

- a. 1 student - 1 examiner  
b. 2 students - 1 examiner  
c. Group of students (3 or more): 1 examiner  
d. No particular preference

Explain your answer

---

---

16) Do you feel that it is appropriate for **students** to veil (partially or fully) for a speaking assessment?

- b. Yes  
c. No  
d. Not sure

Explain your answer

---

---

17) Do you think that it is appropriate for an **examiner** to wear a veil (partially or fully) during the interview?

a. Yes

b. No

c. Not sure

Explain your answer

---

---

Communication during the speaking assessment

18) Do you think that if a student wears a veil, either fully or partially, that this affects the **quality of communication** in the speaking assessment? *Please circle your answer*

a. Yes, it would be better.

b. Yes, it would be worse.

c. No effects

d. Not sure

Explain your answer

---

---

19) Do you think that if a student wears a veil, either fully or partially, that this affects the **grade outcome** in the speaking assessment? *Please circle your answer*

a. Yes, it would be better.

b. Yes, it would be worse.

c. No effects

d. Not sure

Explain your answer

---

---



20) Do you think that the **student** would react positively or negatively to you covering your face, either fully or partially, in the speaking assessment? *Please circle your answer*

- a. Positively   b. Negatively   c. Indifferently (no reaction)   d. Not sure

Explain your answer

---

---

21) Would you prefer to interview online/via telephone/via videoconferencing rather than in a face-to-face interview? *Please circle your answer*

- a. Yes   b. No   c. Not sure

Explain your answer – is gender a factor or are there other reasons affecting your response?

---

---

22) In your experience of teaching English elsewhere in the world, are there any topics that you feel are 'culture sensitive', i.e ones that a particular group of s/s, male or female, would find it difficult to talk about at length, for whatever reason?

d. Yes

e. No

Please give examples

---

---

**SECTION C** (Varieties of English)

23) As an EFL teacher/examiner, can you distinguish between varieties of spoken English, such as British, American and Australian, for example? *Please circle your answer*

- a. Yes      b. No      c. Not sure

Explain your answer

---

---

24) Is it important to you that you speak, **during the exam**, in a particular way? *Please circle your answer*

- a. Yes, \*Standard British    b. Yes, Standard American    c. Yes, Standard Australian    d. No    e. Not sure

Explain your answer

---

---

25) Do you feel that it is important that the student speaks to you, **during the exam**, in a particular way?  
*Please circle your answer*

- a. Yes, British    b. Yes, American    c. Yes, Australian    d. No    e. Not sure

Explain your answer

---

---

\* Standard – accepted academic, non colloquial style.

***THANK YOU SO MUCH FOR ANSWERING THIS QUESTIONNAIRE. ALL INFORMATION GIVEN WILL REMAIN CONFIDENTIAL AND USED ONLY FOR THE PURPOSES OF THIS RESEARCH PROJECT.***

## Appendix 8

### Focus Group transcription sample (excerpt)

Time	Interviewer	Faculty comments
23.06		Another thing is ....there's one question that keeps thick inside
		my head... is in the first part which is supposed to help them relax.
23.14		Not having complete shift in the topics.
23.17		I mean...
	Yeah..	..from..
23.18		Describe your neighborhood.
23.19	Yeah.	
23.2		To talk about the transport....
23.22		Let's talk about
23.25		(hahaha)
23.26	Yeah.	
23.27		It's just like that noise question...
23.28		Noises.
23.31	Yeah there are other strange qs	
23.32		Or even flowers????
23.34		Yeah. Yeah.
23.35		They do make you ask funny qs.
23.36		That I was feel.....
23.38		What the hell am I talking about this for?
23.4	So it doesn't feel natural?.	
23.42		Yeah.
	What other topics?.	
23.45		Mmm.
23.48		vegetables...let's talk about vegetables.
23.5		(haahaha)
23.51		What can you?????by vegetables in your area???
23.53		Yeah...hahahaha
23.56		Sphinx...yes.anytime.
23.59		????those people from Baghdad and I do the one...tell me about
		where you grow up.
24.03		Oh.Yeah...
24.04	Yes, is it a good place to grow up?	It's not a place to grow up.
24.06		And what would you tell this visitors to do there?
24.08		What would you recommend....
24.09		Stay in the green zone????

## Appendix 9

Focus Group transcription sample (excerpt from student focus group)

Interviewer	Students
4.28	Now if you could, what changes would you make to the IELTS
	exam, to the speaking exam that you did.
4.34	Was anything that you would change?
4.38	Maybe, choosing the topic.
4.4	Okay, the topic, why?
4.41	You would like to choose the topic.
4.43	because...
	Yeah. Because sometime you are not er expected to have er er something. Or you are not preparing to answer some questions.
4.52	Mmm.
4.54	So, you feel that it's unfair, maybe, to asked you a particular
	question in the speaking exam.
5	Not unfair but if we have chose the, the topic, it may er maybe we have er more er, er more er more marks
5.09	Mmm.
5.1	Or we have a good average for the speaking exam.
5.12	So you'd, you would want to choose the topics?
5.14	Yes.

## IMAGING SERVICES NORTH

Boston Spa, Wetherby

West Yorkshire, LS23 7BQ

[www.bl.uk](http://www.bl.uk)

**MISSING PAGE/PAGES  
HAVE NO CONTENT**